

5000  
09/486334

REPUBLIQUE FRANÇAISE 09/486334



09/03179  
REC'D 07 JAN 2000  
WIPO PCT  
4

# BREVET D'INVENTION

CERTIFICAT D'UTILITÉ - CERTIFICAT D'ADDITION

COPIE OFFICIELLE

**PRIORITY  
DOCUMENT**  
SUBMITTED OR TRANSMITTED IN  
COMPLIANCE WITH RULE 17.1(a) OR (b)

Le Directeur général de l'Institut national de la propriété industrielle certifie que le document ci-annexé est la copie certifiée conforme d'une demande de titre de propriété industrielle déposée à l'Institut.

Fait à Paris, le 30 NOV. 1999

Pour le Directeur général de l'Institut national de la propriété industrielle  
Le Chef du Département des brevets

Martine PLANCHE

INSTITUT NATIONAL DE LA PROPRIÉTÉ INDUSTRIELLE  
SIEGE  
26 bis, rue de Saint Petersburg  
75800 PARIS Cédex 08  
Téléphone : 01 53 04 53 04  
Télécopie : 01 42 93 59 30





**DIVISION ADMINISTRATIVE DES BREVETS**

26bis, rue de Saint-Petersbourg  
75800 Paris Cédex 08  
Tél. : (1) 42 94 52 52 - Télécopie : (1) 42 93 59 30

N° D'ENREGISTREMENT NATIONAL

2816163

PH 98080

**TITRE DE L'INVENTION :**

Procédé pour augmenter la teneur en cystéine, méthionine et glutathion chez les plantes et plantes obtenues

**LE (S) SOUSSIGNÉ (S)**

RHONE-POULENC AGRO

**DÉSIGNE (NT) EN TANT QU'INVENTEUR (S)** (indiquer nom, prénoms, adresse et souligner le nom patronymique) :

DROUX Michel  
32 Avenue de Lauterbourg  
69169 TASSIN-la-demi-lune (France)

DEROSE Richard  
31 Rue du Bois Guillaume  
91000 EVRY (France)

JOB Dominique  
181 Rue Duguesclin  
69003 LYON (France)

**NOTA** : A titre exceptionnel, le nom de l'inventeur peut être suivi de celui de la société à laquelle il appartient (société d'appartenance) lorsque celle-ci est différente de la société déposante ou titulaire.

Date et signature (s) du (des) demandeur (s) ou du mandataire  
Lyon, le 17 Décembre 1998

TETAZ Franck

**DOCUMENT COMPORTANT DES MODIFICATIONS**

[illegible]

Un changement apporté à la rédaction des revendications d'origine, sauf si celui-ci découle des dispositions de l'article R.612-36 du code de la Propriété Intellectuelle, est signalé par la mention « R.M. » (revendications modifiées).

## Procédé pour augmenter la teneur en cystéine, méthionine et glutathion chez les plantes et plantes obtenues

5 La méthionine est le premier acide aminé essentiel limitant chez les plantes, en particulier les légumineuses qui sont une des bases de l'alimentation animale. La cystéine, autre acide aminé soufré n'est pas un acide aminé essentiel, mais peut être considérée comme un élément limitant pour la nutrition animale puisque la cystéine dérive, chez les animaux, de la méthionine. Dans le maïs, les acides aminés soufrés sont  
10 aussi des acides aminés limitant après la lysine et le tryptophane. En effet les protéines de réserves majoritaires des graines de ces plantes, sont pauvres en ces acides aminés. La surproduction de méthionine et de cystéine dans les graines des légumineuses (soja, luzerne, pois,...) et du maïs aura donc un impact considérable sur la qualité nutritionnelle de ces graines.

15 Jusqu'à présent, l'augmentation de la qualité nutritionnelle des aliments dérivés des graines de légumineuses a été obtenue par complémentation avec de la méthionine libre synthétisée chimiquement. Par exemple, les contenus moyens en méthionine + cystéine du soja et du pois sont de l'ordre de 20 mg par g de protéines. Ce contenu doit être augmenté à une valeur de l'ordre de 25 mg cystéine + méthionine/g de protéines  
20 pour couvrir les besoins alimentaires de l'homme adulte, et à une valeur de l'ordre de 48 mg de cystéine + méthionine/g de protéines pour couvrir ceux des porcs (De Lumen, B.O., Food Technology (1997) 51, 67-70).

Les techniques de caractérisation des protéines enrichies en acides aminés soufrés et la préparation de plantes transgéniques permettant l'expression de telles  
25 protéines, de manière à augmenter la teneur en acides aminés soufrés de ces plantes, et donc leur valeur nutritive pour l'alimentation animale, donc à diminuer l'apport en méthionine de synthèse, sont maintenant bien connus et décrites dans la littérature([1] Korit, A.A. & al., Eur. J. Biochem. (1991) 195, 329-334 ; WO 98/20133 ; WO 97/41239 ; WO 95/31554 ; WO 94/20828 ; WO 92/14822).

30 L'enrichissement en protéines à fortes teneurs en acides aminés soufrés par une telle approche reste toutefois limité à la capacité des cellules végétales et des plantes à produire lesdits acides aminés soufrés. Or, on a constaté que le paramètre limitant d'une telle approche est bien lié à cette capacité à produire de la méthionine. Il est donc important de pouvoir modifier dans les plantes cette capacité à produire de la  
35 méthionine en quantités suffisantes pour permettre la production de protéines hétérologues à haute teneur en acides aminés soufrés, c'est à dire de mettre en œuvre une stratégie moléculaire visant à augmenter les taux de cystéine et de méthionine chez les plantes, et plus particulièrement les plantes de cultures d'intérêt agronomique.

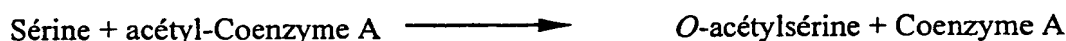
Chez les plantes, on connaît que la biosynthèse de méthionine est effectuée à partir de la cystéine, cette même cystéine étant impliquée dans la synthèse du glutathion.

Le glutathion est une forme de stockage du soufre réduit et représente 60 à 70% du soufre organique dans la cellule. Le glutathion joue un rôle important pour les plantes dans la résistance contre le stress oxydatif et l'élimination des composés toxiques. Il participe ainsi à l'élimination des composés xénobiotiques: des métaux lourds (par exemple) via la formation des phytochélatines; des herbicides, via l'activité glutathion S-transférase; qui sont toxiques pour la plante, et dans des mécanismes de défenses de la plante contre les micro-organismes. En augmentant la teneur en cystéine d'une plante, et par conséquent sa teneur en glutathion, il est alors possible de moduler la réponse de la plante aux différents stress cités ci-dessus.

Il existe donc à partir de la cystéine, deux voies métaboliques distinctes, l'une pour la préparation de la méthionine, l'autre pour la préparation du glutathion, dont les différentes enzymes impliquées sont rappelées ci-après.

Chez les plantes les étapes finales de synthèse de la cystéine impliquent les deux enzymes suivantes :

**E1) Sérine acétyltransférase (EC 2.3.1.30) (SAT):**

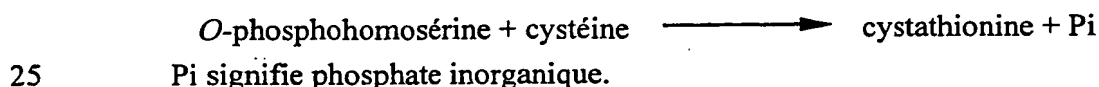


**E2) O-acétylsérine (thiol) lyase (EC 4.2.99.8) (OASTL):**



La synthèse de la méthionine à partir de la cystéine implique une succession des trois enzymes suivantes :

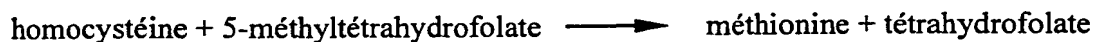
**E3) cystathionine  $\gamma$ -synthase (EC 4.2.99.9) (CGS) :**



**E4) cystathionine  $\beta$ -lyase (EC 4.4.1.8) (CBL) :**

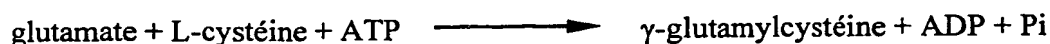


**E5) méthionine synthase (EC 2.1.1.14) (MS) :**

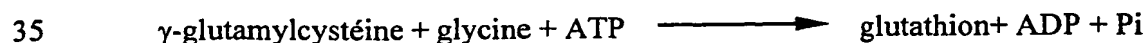


La synthèse du glutathion à partir de la cystéine implique pour sa part une succession des deux enzymes suivantes :

**E6)  $\gamma$ -glutamylcystéine synthétase (EC 6.3.2.2)**



**E7) glutathion synthétase (EC 6. 3.2.3)**



Toutes ces enzymes ont été caractérisées et clonées chez les plantes ([2] Lunn, J.E. & al., *Plant Physiol.* (1990) 94, 1345-1352 ; [3] Rolland, N & al., *Plant Physiol.* (1992) 98, 927-935 ; [4] Droux, M. & al., *Arch. Biochem. Biophys.* (1992) 295, 379-390 ; [5] Rolland, N. & al., *Arch. Biochem.* (1993) 300, 213-222 ; [6] Ruffet, M.L. & al., *Plant Physiol.* (1994) 104, 597 - 604 ; [7] Ravanel, S. & al., *Arch. Biochem. Biophys.* (1995) 316, 572 - 5584 ; [8] Droux, M. & al., *Arch. Biochem. Biophys.* (1995) 31, 585 - 595 ; [9] Ruffet, M.L. & al., *Eur. J. Biochem.* (1995) 227, 500 - 509 ; [10] Ravanel, S. & al., *Biochem. J.* (1996) 320, 383 - 392 ; [11] Ravanel, S. & al., *Plant Mol. Biol.* (1996) 29, 875 - 882 ; [12] Rolland, N. & al., *Eur. J. Biochem.* (1996) 236, 272 - 282 ; [13] Ravanel, S. & al., *Biochem. J.* (1998) 331, 639-648 ; [14] Droux, M & al., *Eur. J. Biochem.* (1998) 255, 235-245 ; [15] May, M.J., Leaver, C.J., *Proc. Natl. Acad. Sci. USA* (1994) 91, 10059-10063 ; [16] Ullmann, P. & al., *Eur. J. Biochem.* (1996) 236, 662-669 ; [17] Eichel, J. & al., *Eur. J. Biochem.* (1995) 230, 1053-1058).

On sait que pour la synthèse de cystéine, les enzymes E1 et E2 sont présentes dans les trois compartiments de la cellule végétale, c'est-à-dire, les plastes, le cytosol et les mitochondries (2-3, 6, 9). Ces trois enzymes E1 sont dénomées SAT1 pour l'enzyme chloroplastique, SAT2 pour l'enzyme mitochondriale et SAT3 pour l'enzyme cytoplasmique.

Pour les enzymes de la synthèse de la méthionine, la situation est différente puisque les enzymes E3 et E4 sont exclusivement localisées dans les plastes (7-8, 10-11, 13), alors que l'enzyme terminale E5 est dans le cytosol (17).

Les enzymes associées à la voie de biosynthèse du glutathion sont pour leur part localisées à la fois dans le chloroplaste et le cytosol ([18] Hell, R. and Bergmann, L., *Planta* (1990) 180, 603-612).

L'enzyme E3, de la voie de synthèse de la méthionine, présente un  $K_m$  (concentration en substrat donnant la moitié de la vitesse maximum) de l'ordre de 200  $\mu M$  à 500  $\mu M$  pour la cystéine (7, 13, [19] Kreft, B-D. & al., *Plant Physiol.* (1994) 104, 1215-1220).

L'enzyme E6, de la voie de synthèse du glutathion, présente aussi un  $K_m$  élevé pour la cystéine, de l'ordre de 200  $\mu M$  [18].

On a maintenant constaté que les enzymes SAT1 chloroplastiques (figure 1) était inhibée par la cystéine, contrairement à l'enzyme SAT3 cytoplasmique (figure 1), cette inhibition constituant le facteur limitant essentiel de la synthèse de cystéine dans les cellules végétales, et en aval de la méthionine et du glutathion.

La présente invention consiste donc à augmenter le taux de cystéine synthétisée dans le compartiment chloroplastique de manière à augmenter le taux de cystéine produite par les cellules végétales et les plantes. L'augmentation du taux de ce précurseur soufré (la cystéine) permet avantageusement d'augmenter le taux de



méthionine et/ou de glutathion dans les cellules végétales et les plantes, et subséquemment à améliorer la production de protéines, naturelles ou hétérologues, enrichies en acides aminés soufrés dans les cellules végétales et les plantes.

Cette augmentation selon l'invention est obtenue en surexprimant une Sérine  
5 acétyltransférase (SAT) dans les chloroplastes.

La présente invention concerne donc un procédé pour augmenter la production de cystéine par les cellules végétales et les plantes, ledit procédé consistant à surexprimer une SAT dans les chloroplastes des cellules végétales, et des plantes contenant les dites cellules végétales.

10 La SAT surexprimée dans les chloroplastes peut être constituée par toute SAT, qu'elle soit d'origine végétale, notamment SAT1, SAT2 ou SAT3, ou de toute autre origine, notamment bactérienne, sous une forme native ou mutée.

Elle peut notamment être une SAT sensible à la cystéine, comme par exemple une SAT chloroplastique de plante (SAT1), ou une SAT native d'origine bactérienne  
15 (Nakamori et al., 1998, Appl. Environ. Microbiol., 64, 1607-1611.).

Selon un autre mode de réalisation de l'invention elle peut être une SAT insensible à la cystéine, comme par exemple une SAT cytoplasmique de plante (SAT3), ou une SAT d'origine bactérienne mutée, rendue insensible à la cystéine par mutagenèse (Nakamori et al., 1998, Appl. Environ. Microbiol., 64, 1607-1611 dont le contenu est  
20 incorporé ici par référence), ou toutes SAT de plantes mutée et fonctionnelle dans la synthèse de l'O-acétylsérine (le précurseur carbonné pour la synthèse de la cystéine)

Selon un mode particulier de réalisation de l'invention, la SAT est une SAT d'*Arabidopsis thaliana* (9).

De manière avantageuse, la protéine SAT est choisie parmi les protéines SAT3  
25 (L34076), SAT3' (U30298) et SAT1 (L78443 ou U22964) représentées avec leur séquence codante par les identificateurs de séquence 1, 2, et 3 et 4 respectivement (SEQ ID NO 1, SEQ ID NO 2, SEQ ID NO 3 et SEQ ID NO 4 en annexe).

La SAT sera exprimée dans les chloroplastes par tout moyen approprié, en particulier par tout moyen connu de l'homme du métier et largement décrit dans l'état  
30 de la technique.

Selon un premier mode de réalisation de l'invention, la SAT est surexprimée dans les chloroplastes par intégration dans l'ADN chloroplastique d'un gène chimère comprenant une séquence d'ADN codant pour ladite SAT, sous le contrôle d'éléments de régulations en 5' et 3' fonctionnels dans les chloroplastes. Les techniques d'insertion  
35 de gènes dans les chloroplastes, comme les éléments de régulations appropriés à l'expression desdits gènes dans les chloroplastes sont bien connus de l'homme du métier, et notamment décrits dans les brevets et demandes de brevet suivants : US 5,693,507, US 5,451,513 et WO 97/32977.

Selon un autre mode particulier de réalisation de l'invention, la SAT est surexprimée dans le cytoplasme sous la forme d'une protéine de fusion peptide de transit/SAT, le peptide de transit ayant pour fonction d'adresser la SAT à laquelle il est fusionné vers les chloroplastes, la SAT mature fonctionnelle étant libérée à l'intérieur  
5 des chloroplastes après clivage au niveau de la membrane chloroplastique.

Les peptides de transit, leurs structures, leurs modes de fonctionnement et leur utilisation dans la construction de gènes chimères pour l'adressage d'une protéine hétérologue dans les chloroplastes, ainsi que des peptides de transit chimères comprenant plusieurs peptides de transit sont bien connus de l'homme du métier et  
10 largement décrits dans la littérature. On citera notamment les demandes de brevet suivants : EP 189 707, EP 218 571 et EP 508 909, et les références citées dans ces demandes de brevet, dont le contenu est incorporé ici par référence.

La présente invention concerne donc une protéine de fusion peptide de transit/SAT, dans laquelle la SAT est définie plus haut, et dans laquelle le peptide de  
15 transit est constitué par au moins un peptide de transit d'une protéine végétale naturelle à localisation plastidiale. On entend par protéine à localisation plastidiale une protéine exprimée dans le cytoplasme des cellules végétales sous la forme d'une protéine de fusion peptide de transit/protéine, la protéine mature étant localisée dans le chloroplaste après clivage du peptide de transit.

20 Dans la protéine de fusion selon l'invention, la SAT peut être homologue ou hétérologue du peptide de transit. Dans le premier cas, la protéine de fusion est la protéine SAT1 exprimée naturellement dans les cellules végétales. Dans le second cas, le peptide de transit peut être un peptide de transit d'une SAT1 d'une autre espèce végétale, ou un peptide de transit d'une autre protéine à localisation plastidiale.

25 Il est entendu que dans le cas de la protéine de fusion hétérologue, la SAT peut être constituée par toute SAT, qu'elle soit d'origine végétale SAT1, SAT2 ou SAT3, ou de toute autre origine, notamment bactérienne, en particulier les SAT définies précédemment.

Un peptide de transit d'EPSPS de plante est notamment décrit dans la demande  
30 de brevet EP 218 571, alors qu'un peptide de transit de ssu RuBisCO de plante est décrit dans la demande de brevet EP 189 707.

Selon un autre mode de réalisation de l'invention, le peptide de transit comprend également, entre la partie C-terminale du peptide de transit et la partie N-terminale de la SAT une partie de séquence de la partie mature N-terminale d'une protéine à localisation  
35 plastidiale, cette partie de séquence comprenant généralement moins de 40 acides aminés de la partie N-terminale de la protéine mature, de préférence moins de 30 acides aminés, plus préférentiellement entre 15 et 25 acides aminés. Un tel peptide de transit comprenant un peptide de transit fusionné à une partie de la partie N-terminale d'une protéine à localisation plastidiale est notamment décrit dans la demande de brevet EP

189 707, plus particulièrement pour le peptide de transit et la partie N-terminale de ssu RuBisCO de plante.

Selon un autre mode de réalisation de l'invention, le peptide de transit comprend également entre la partie C-terminale de la partie N-terminale de la protéine mature et la  
 5 partie N-terminale de la SAT un deuxième peptide de transit d'une protéine végétale à localisation plastidiale. Préférentiellement, ce peptide de transit chimère comprenant l'association de plusieurs peptide de transit est un peptide de transit optimisé (OTP) constitué par la fusion d'un premier peptide de transit, avec une partie de séquence de la  
 10 partie mature N-terminale d'une protéine à localisation plastidiale, laquelle est fusionnée avec un deuxième peptide de transit. Un tel peptide de transit optimisé est décrit dans la demande de brevet EP 508 909, plus particulièrement le peptide de transit optimisé comprenant le peptide de transit de la ssu RuBisCO de tournesol, fusionné à un peptide constitué par les 22 acides aminés de la partie N-terminale de la ssu RuBisCO de maïs  
 mature, fusionné au peptide de transit de la ssu RuBisCO de maïs.

15 La présente invention concerne également une séquence d'acide nucléique codant pour une protéine de fusion peptide de transit/SAT décrite ci-dessus. Selon la présente invention, on entend par « séquence d'acide nucléique » une séquence nucléotidique pouvant être de type ADN ou ARN, de préférence de type ADN, notamment double brin, qu'elle soit d'origine naturelle ou synthétique, notamment une  
 20 séquence d'ADN pour laquelle les codons codant pour la protéine de fusion selon l'invention auront été optimisés en fonction de l'organisme hôte dans lequel elle sera exprimée, ces méthodes d'optimisations étant bien connues de l'homme du métier.

L'invention a encore pour objet l'utilisation d'une séquence d'acide nucléique codant pour une protéine de fusion selon l'invention dans un procédé pour la  
 25 transformation des plantes, comme séquence codante permettant de modifier le contenu en cystéine, méthionine et glutathion des plantes transformées. Cette séquence peut bien entendu également être utilisée en association avec d'autre(s) gène(s) marqueur(s) et/ou séquence(s) codante(s) pour une ou plusieurs autres propriétés agronomiques.

La présente invention concerne également un gène chimère (ou cassette  
 30 d'expression) comprenant une séquence codante ainsi que des éléments de régulation en position 5' et 3' hétérologues pouvant fonctionner dans un organisme hôte, en particulier les cellules végétales ou les plantes, la séquence codante comprenant au moins une séquence d'acide nucléique codant pour une protéine de fusion telle que définie précédemment.

35 Par organisme hôte, on entend tout organisme mono ou pluricellulaire, inférieur ou supérieur, dans lequel le gène chimère selon l'invention peut être introduit, pour la production de protéine de fusion peptide de transit/SAT. Il s'agit en particulier de bactéries, par exemple *E. coli*, de levures, en particulier des genres *Saccharomyces* ou *Kluyveromyces*, *Pichia*, de champignons, en particulier *Aspergillus*, d'un baculovirus,

ou de préférence des cellules végétales et des plantes.

Par "cellule végétale", on entend selon l'invention toute cellule issue d'une plante et pouvant constituer des tissus indifférenciés tels que des cals, des tissus différenciés tels que des embryons, des parties de plantes, des plantes ou des semences.

5 On entend par "plante" selon l'invention, tout organisme multicellulaire différencié capable de photosynthèse, en particulier monocotylédones ou dicotylédones, plus particulièrement des plantes de culture destinées ou non à l'alimentation animale ou humaine, comme le maïs, le blé, le colza, le soja, le riz, la canne à sucre, la betterave, le tabac, le coton, etc.

10 Les éléments de régulation nécessaires à l'expression de la séquence d'acide nucléique codant pour une protéine de fusion selon l'invention sont bien connus de l'homme du métier en fonction de l'organisme hôte. Ils comprennent notamment des séquences promotrices, des activateurs de transcription, des séquences terminatrices, y compris des codons start et stop. Les moyens et méthodes pour identifier et sélectionner  
15 les éléments de régulation sont bien connus de l'homme du métier et largement décrits dans la littérature.

L'invention concerne plus particulièrement la transformation des plantes. Comme séquence de régulation promotrice dans les plantes, on peut utiliser toute séquence promotrice d'un gène s'exprimant naturellement dans les plantes en particulier  
20 un promoteur s'exprimant notamment dans les feuilles des plantes, comme par exemple des promoteurs dits constitutifs d'origine bactérienne, virale ou végétale ou encore des promoteurs dits lumière dépendants comme celui d'un gène de la petite sous-unité de ribulose- biscarboxylase/oxygénase (RuBisCO) de plante ou tout promoteur convenable connu pouvant être utilisé. Parmi les promoteurs d'origine végétale on citera les  
25 promoteurs d'histone tels que décrits dans la demande EP 0 507 698, ou le promoteur d'actine de riz (US 5,641,876). Parmi les promoteurs d'un gène de virus de plante, on citera celui de la mosaïque du choux fleur (CAMV 19S ou 35S), ou le promoteur du circovirus (AU 689 311).

On peut encore utiliser une séquence de régulation promotrice spécifique de  
30 régions ou de tissus particuliers des plantes, et plus particulièrement des promoteurs spécifiques des graines ([22] Datla, R.& al., Biotechnology Ann. Rev. (1997) 3, 269-296), notamment les promoteurs de la napine (EP 255 378), de la phaseoline, de la glutenine, de l'héliantinine (WO 92/17580), de l'albumine (WO 98/45460), de l'oélosine (WO 98/45461), de l'ATS1 ou de l'ATS3 (PCT/US98/06978, déposée le 20  
35 octobre 1998, incorporée ici par référence).

Selon l'invention, on peut également utiliser, en association avec la séquence de régulation promotrice, d'autres séquences de régulation, qui sont situées entre le promoteur et la séquence codante, telles que des activateurs de transcription ("enhancer"), comme par exemple l'activateur de translation du virus de la mosaïque du

tabac (TMV) décrit dans la demande WO 87/07644, ou du virus etch du tabac (TEV) décrit par Carrington & Freed.

Comme séquence de régulation terminatrice ou de polyadénylation, on peut utiliser toute séquence correspondante d'origine bactérienne, comme par exemple le  
 5 terminateur nos d'*Agrobacterium tumefaciens*, ou encore d'origine végétale, comme par exemple un terminateur d'histone tel que décrit dans la demande EP 0 633 317.

La présente invention concerne également un vecteur de clonage et/ou d'expression pour la transformation d'un organisme hôte contenant au moins un gène chimère tel que défini ci-dessus. Ce vecteur comprend outre le gène chimère ci-dessus,  
 10 au moins une origine de répllication. Ce vecteur peut être constitué par un plasmide, un cosmide, un bactériophage ou un virus, transformés par l'introduction du gène chimère selon l'invention. De tels vecteurs de transformation en fonction de l'organisme hôte à transformer sont bien connus de l'homme du métier et largement décrits dans la littérature. Pour la transformation des cellules végétales ou des plantes, il s'agira  
 15 notamment d'un virus qui peut être employé pour la transformation des plantes développées et contenant en outre ses propres éléments de répllication et d'expression. De manière préférentielle, le vecteur de transformation des cellules végétales ou des plantes selon l'invention est un plasmide.

Pour la transformation des organismes hôtes, le gène chimère selon l'invention  
 20 peut être employé en association avec un gène marqueur de sélection, soit dans un même vecteur, les deux gènes étant associés de manière convergente, divergente ou colinéaire, ou encore dans deux vecteurs employés simultanément pour la transformation de l'organisme hôte. De tels gènes marqueurs et leur utilisation pour la transformation d'organismes hôtes sont bien connus de l'homme du métier et largement  
 25 décrits dans la littérature.

Parmi les gènes codant pour des marqueurs de sélection, on peut citer les gènes de résistance aux antibiotiques, les gènes de tolérance aux herbicides (bialaphos, glyphosate ou isoxazoles), des gènes codant pour des enzymes facilement identifiables comme l'enzyme GUS, des gènes codant pour des pigments ou des enzymes régulant la  
 30 production de pigments dans les cellules transformées. De tels gènes marqueurs de sélection sont notamment décrits dans les demandes de brevet EP 242 236, EP 242 246, GB 2 197 653, WO 91/02071, WO 95/06128, WO 96/38567 ou WO 97/04103.

L'invention a encore pour objet un procédé de transformation des organismes hôtes, en particulier des cellules végétales par intégration d'au moins une séquence  
 35 d'acide nucléique ou un gène chimère tels que définis ci-dessus, transformation qui peut être obtenue par tout moyen connu approprié, amplement décrit dans la littérature spécialisée et notamment les références citées dans la présente demande, plus particulièrement par le vecteur selon l'invention.

Une série de méthodes consiste à bombarder des cellules, des protoplastes ou des tissus avec des particules auxquelles sont accrochées les séquences d'ADN. Une autre série de méthodes consiste à utiliser comme moyen de transfert dans la plante un gène chimère inséré dans un plasmide Ti d'*Agrobacterium tumefaciens* ou Ri d'*Agrobacterium rhizogenes*. D'autres méthodes peuvent être utilisées telles que la micro-injection ou l'électroporation, ou encore la précipitation directe au moyen de PEG. L'homme du métier fera le choix de la méthode appropriée en fonction de la nature de l'organisme hôte, en particulier de la cellule végétale ou de la plante.

La présente invention a encore pour objet les organismes hôtes, en particulier cellules végétales ou plantes, transformés et contenant un gène chimère défini ci-dessus.

La présente invention a encore pour objet les plantes contenant des cellules transformées, en particulier les plantes régénérées à partir des cellules transformées. La régénération est obtenue par tout procédé approprié qui dépend de la nature de l'espèce, comme par exemple décrit dans les références ci-dessus. Pour les procédés de transformation des cellules végétales et de régénération des plantes, on citera notamment les brevets et demandes de brevet suivants: US 4,459,355, US 4,536,475, US 5,464,763, US 5,177,010, US 5,187,073, EP 267,159, EP 604 662, EP 672 752, US 4,945,050, US 5,036,006, US 5,100,792, US 5,371,014, US 5,478,744, US 5,179,022, US 5,565,346, US 5,484,956, US 5,508,468, US 5,538,877, US 5,554,798, US 5,489,520, US 5,510,318, US 5,204,253, US 5,405,765, EP 442 174, EP 486 233, EP 486 234, EP 539 563, EP 674 725, WO 91/02071 et WO 95/06128.

La présente invention concerne également les plantes transformées issues de la culture et/ou du croisement des plantes régénérées ci-dessus, ainsi que les graines de plantes transformées.

Les plantes transformées pouvant être obtenues selon l'invention peuvent être du type monocotylédones telles que par exemple les céréales, la canne à sucre, le riz et le maïs ou dicotylédones comme par exemple le tabac, la soja, le colza, le coton, la betterave, le trèfle, etc.

Les plantes transformées selon l'invention peuvent contenir d'autres gènes d'intérêt, conférant aux plantes de nouvelles propriétés agronomiques. Parmi les gènes conférant de nouvelles propriétés agronomiques aux plantes transformées, on peut citer les gènes conférant une tolérance à certains herbicides, ceux conférant une tolérance à certains insectes, ceux conférant une tolérance à certaines maladies. De tels gènes sont notamment décrits dans les demandes de brevet WO 91/02071 et WO 95/06128. On peut également citer les gènes modifiant la constitution des plantes modifiées, en particulier la teneur et la qualité de certains acides gras essentiels (EP 666 918) ou encore la teneur et la qualité des protéines, en particuliers dans les feuilles et/ou les graines desdites plantes. On citera en particulier les gènes codant pour des protéines enrichies en acides aminés soufrés ([1]; WO 98/20133 ; WO 97/41239 ; WO 95/31554 ;

WO 94/20828 ; WO 92/14822). Ces protéines enrichies en acides aminés soufrés auront également pour fonction de piéger et stocker la cystéine et/ou la méthionine excédentaire, permettant d'éviter les problèmes éventuels de toxicité liés à une surproduction de ces acides aminés soufrés en les piégeant.

- 5 On peut citer également des gènes codant pour des peptides riches en acides aminés soufrés et plus particulièrement en cystéines, les dits peptides ayant également une activité antibactérienne et/ou antifongique. On citera plus particulièrement les défensines de plantes, de même que les peptides lytiques de toute origine, et plus particulièrement les peptides lytiques suivants : l'androctonine (WO 97/30082 et  
10 PCT/FR98/01814, déposée le 18 août 1998), la drosomicine (PCT/FR98/01462, déposée le 8 juillet 1998), la thanatine (PCT/FR98/02375, déposée le 6 novembre 1998) ou l'héliomicine (FR 98 04933, déposée le 15 avril 1998).

- Ces autres gènes d'intérêt peuvent être combinés au gène chimère selon l'invention soit par croisement conventionnel de deux plantes contenant chacune l'un  
15 des gènes (la première le gène chimère selon l'invention et la seconde le gène codant pour la protéine d'intérêt) soit en effectuant la transformation de cellules végétales d'une plante contenant le gène codant pour la protéine d'intérêt avec le gène chimère selon l'invention.

- Les exemples ci-après permettent d'illustrer l'invention, sans toutefois chercher  
20 à en limiter la portée.

- Toutes les méthodes ou opérations décrites ci-dessous dans ces exemples sont données à titre d'exemples et correspondent à un choix, effectué parmi les différentes méthodes disponibles pour parvenir au même résultat. Ce choix n'a aucune incidence sur la qualité du résultat et par conséquent, toute méthode adaptée peut être utilisée par  
25 l'homme de l'art pour parvenir au même résultat. La plupart des méthodes d'ingénierie des fragments d'ADN sont décrites dans "Current Protocols in Molecular Biology" Volumes 1 et 2, Ausubel F.M. et al , publiés par Greene Publishing Associates et Wiley -Interscience (1989) ou dans Molecular cloning, T.Maniatis, E.F.Fritsch, J.Sambrook, 1982.

30

**Exemple 1. Mise en évidence de l'inhibition de la serine acetyltransferase chloroplastique de feuilles de pois (*Pisum sativum*) par la cystéine**

- On prépare à partir de feuilles de pois les trois compartiments subcellulaires correspondant au cytosol (préparation à partir d'un fractionnement subcellulaire de  
35 protoplastes de pois, [9]), aux mitochondries et aux chloroplastes [9]. On en extrait les protéines solubles et on mesure l'activité sérine acétyltransférase au moyen d'une technique décrite [9, 14].

Description de la méthode de dosage:

L'activité de la sérine acétyltransférase est mesurée par chromatographie liquide haute performance (HPLC), en dosant l'*O*-acétylsérine formée au cours de la réaction (réaction 1), après dérivatisation avec de l'orthophtalaldéhyde (OPA). Une quantité définie de l'extrait protéique, correspondant au cytosol, et aux différentes fractions solubles des chloroplastes (stroma) et mitochondries (matrice), sont dessalées sur une colonne PD10 (Pharmacia) préalablement équilibrée en tampon 50 mM Hepes-HCl, pH 7,5 et 1 mM EDTA. La réaction enzymatique est effectuée en présence de 50 mM Hepes-HCl, pH 7,5, 1mM dithiotreitol, 10 mM L-sérine, 2,5 mM acétyl-CoA, dans un volume réactionnel de 100 µl, à 25 °C. Après 10 à 15 min d'incubation, la réaction est arrêtée par l'addition de 50 µl d'acide trichloroacétique 20% (P/V). Les protéines précipitées sont ensuite éliminées par une centrifugation de 2 min à 15,000g. Le surnageant, contenant le produit de la réaction, est mélangé avec 500 µl d'une solution de dérivatisation (54 mg d'OPA dissous dans 1 ml d'éthanol absolu, 9 ml d'une solution de borate-NaOH 400 mM, pH 9,5 et 0,2 ml de β-mercaptoethanol 14 M) et incubée pendant 2 min. Une fraction de ce mélange (20 µl) est injectée sur une colonne de phase inverse (colonne AccQ Tag C18, 3,9 X 150 mm, Waters) connectée à un système HPLC. Les tampons utilisés pour l'élution des composés dérivés par l'OPA sont: Tampon A, 85 mM acétate de sodium, pH 4,5, et acétonitrile 6% (V/V), pH 4,5; Tampon B, acétonitrile, 60 % (V/V) dans de l'eau. L'*O*-acétylsérine, dérivée par l'OPA, est éluée par un gradient linéaire continu du tampon B dans le tampon A de 25 à 70 % (V/V), et détectée par fluorescence à 455 nm (excitation à 340 nm). Le temps de rétention de l'*O*-acétylsérine est de l'ordre de 6.2 min, et la quantité de produit formée dans les essais enzymatiques est quantifiée à partir d'une courbe étalon réalisée avec de l'*O*-acétylsérine. Les essais enzymatiques ont été optimisés afin de respecter le pH optimal de la réaction, la linéarité en fonction du temps, et afin d'opérer dans des concentrations saturantes en substrats.

#### Effet de la cystéine sur l'activité sérine acétyltransférase de feuilles de pois

L'incubation (2 min) est réalisée en présence de l'extrait protéique (cytosol, matrice et stroma), et des concentrations croissantes de L-cystéine (de 0 à 1 mM), avant l'ajout de concentrations saturantes des substrats de la sérine acétyltransférase, la L-sérine (10 mM) et l'acétyl-CoA (2,5 mM). La réaction enzymatique et le dosage de l'*O*-acétylsérine résiduelle dans le surnageant sont effectués comme décrit précédemment. Le résultat de ces expériences est représenté sur le graphe de la figure 1 en annexe.

L'activité sérine acétyltransférase résiduelle associée à chacun des compartiments cellulaires (le chloroplaste, la mitochondrie et le cytosol) est mesurée après une incubation de l'extrait en présence de concentrations croissantes de L-cystéine.

Si on ajoute de la cystéine libre (de 0 à 1 mM, figure 1) aux différents essais, on constate une très forte inhibition de l'activité sérine acétyltransférase chloroplastique. L'activité sérine acétyltransférase mitochondriale est inhibée mais pour des



concentrations supérieures de cystéine (constante d'inhibition de l'ordre de 300  $\mu\text{M}$ ). En revanche, l'activité sérine acétyltransférase cytosolique est insensible à l'inhibition par la cystéine jusqu'à des concentrations de l'ordre de 1 mM (figure 1). Ce résultat prouve donc que seul l'activité sérine acétyltransférase chloroplastique, donc l'enzyme associée à la voie de l'assimilation du sulfate, est inhibée par le produit finale, la L-cystéine.

**Tableau I: Détermination des activité spécifiques et des valeurs de  $\text{CI}_{50}$  pour la cystéine pour chacune des isoformes de la serine acétyltransférase.**

<b>Serine acétyltransférase (<i>Pisum sativum</i>)</b>		
	<b>Activité spécifique</b> <b>nmoles OAS <math>\cdot \text{min}^{-1} \cdot \text{mg}^{-1}</math></b>	<b><math>\text{CI}_{50}</math> L-cystéine</b> <b><math>\mu\text{M}</math></b>
<b>Stroma</b>	$0,93 \pm 0,2$	$33,4 \pm 8$
<b>Matrice</b>	$10 \pm 2$	$283 \pm 50$
<b>Cytosol</b>	$0,83 \pm 0,3$	pas d'inhibition

La concentration en L-cystéine permettant l'obtention de 50% d'inhibition ( $\text{CI}_{50}$ ) dans les conditions standards de la réaction calculée pour différentes préparations enzymatiques est représenté dans le tableau I. Les déterminations de l'activité enzymatique de la serine acétyltransférase et de la  $\text{CI}_{50}$  ont été effectuées lors de 9 expériences différentes (stroma), et 3 pour les extraits cytosolique et issus de la mitochondrie.

#### Etude du mode d'inhibition de l'activité serine acétyltransférase par la cystéine

La vitesse de la réaction enzymatique a été déterminée pour des concentrations fixes de cystéine (0  $\mu\text{M}$ ; 10  $\mu\text{M}$ ; 20  $\mu\text{M}$ ; 40  $\mu\text{M}$ ; 60  $\mu\text{M}$  et 100  $\mu\text{M}$ ) en faisant varier soit la concentration en L-sérine ou en acétyl-CoA, pour des concentrations saturantes du second co-substrat. Pour chacune des cinétiques obtenues, l'affinité de l'enzyme pour ces substrats ne semblent pas être affectée, mais par contre la vitesse maximale de la réaction est modifiée. Plus la concentration en L-cystéine augmente, plus la vitesse de synthèse de l'O-acétylsérine est diminuée. Pour chacune des conditions analysées, la constante d'inhibition  $K_i$  a été estimée de l'ordre de 30 ( $\pm 2,2$ )  $\mu\text{M}$  (substrat variable: la sérine), et de 22 ( $\pm 2$ )  $\mu\text{M}$  (substrats variable: l'acétyl-CoA). Nous avons pu montré que la cystéine est un inhibiteur de type non-compétitif pour l'activité sérine acétyltransférase et de plus de type allostérique (constante de Hill de l'ordre de  $1,6 \pm 0,3$   $\mu\text{M}$ ) en utilisant les équations cinétiques classiques ([20] Segel, I.H. (1995), John Wiley and Sons, New-York). Ces résultats indiquent que l'inhibition de l'enzyme chloroplastique prend place en un site différent du site actif, et de plus qui n'existe pas sur l'isoforme sérine acétyltransférase, présent dans le cytosol.

Ces valeurs de constantes d'inhibition sont consistantes avec la concentration en cystéine déterminée dans les chloroplastes de pois de  $40 \pm 10 \mu\text{M}$  (2 nmoles / mg chlorophylle), valeur qui est calculée pour un volume du compartiment stromatique de l'ordre de 35 à 65  $\mu\text{l}$  par mg de chlorophylle.

5

**Exemple 2. Isolement et caractérisation d'un gène codant pour une isoforme serine acétyltransférase putative cytoplasmique (SAT3) [9]**

On reprend dans cet exemple le mode opératoire décrit en page 502 de Ruffet & al [9], en particulier les chapitres décrits sous les titres « Bacterial strain and growth conditions » et « Isolation of *A. thaliana* serine acetyltransferase cDNA clones by complementation in *E. coli* ».

Un gène codant pour une serine acétyltransférase putative cytosolique (Z34888 ou L34076) représentée sur la figure 4 (SEQ ID NO 1), a été isolé par complémentation fonctionnelle d'une souche d'*Escherichia coli* déficiente pour l'activité serine acétyltransférase. L'analyse de la séquence primaire a montré la présence d'une forte similitude avec la séquence de l'enzyme bactérienne (40% d'identité).

Les amorces qui ont été utilisées pour l'amplification de la séquence nucléotidique et son clonage dans le vecteur utilisé pour la transformation de plants de tabac (en caractères gras sur la figure 2) sont les suivantes :

20 Oligo 1 : 5' GAGAGAGGAT CCTCTTTCCA ATCATAAACC ATGGCAACAT  
GCATAGACAC ATGC 3'  
Oligo 2 : 5' GGCTCACCAG ACTAATACAC TAAATTGTGT TTACCTCGAG  
AGAGAG 3'

Ces amorces permettent l'introduction des sites de restriction *Bam*HI en 5' (GGATCC) et *Sac*I en 3' (GAGCTC).

L'extrémité N-terminale de la séquence en acide aminé de l'isoforme SAT3 ne présente pas les caractéristiques des peptides d'adressage dans un organite (mitochondrie ou chloroplaste). Cette analyse conduit à supposer une localisation cytosolique pour cette isoforme [9].

30

**Exemple 3. Sur-expression et purification de la SAT3 chez *Escherichia coli***

Le protocole défini pour la sur-expression de l'enzyme chez *E. coli* permet la purification de l'enzyme sous sa forme libre ou en complexe avec l'O-acetylserine (thiol) lyase de plante, le complexe cystéine synthase [14]. Avec les protéines purifiées, l'effet de la cystéine sur l'activité de la serine acétyltransférase a été analysée par un dosage spectrophotométrique basé sur la consommation de l'acétyl-CoA au cours de la réaction 1, en fonction du temps d'incubation. Cette analyse s'effectue dans un milieu (1 ml) contenant 50 mM Hepes-HCl, pH 7.5, 2 mM L-serine et 0.2 mM Acetyl-CoA. La

réaction est suivie en mesurant la diminution de l'absorbance à 232 nm (coefficient d'extinction moléculaire de  $4200 \text{ M}^{-1}\text{cm}^{-1}$ ) ([21] Kredich, N.M. & al., J. Biol. Chem. (1969) 244, 2428-2439). Nous avons pu montrer que cette isoforme (SAT3) sous sa forme libre ou en complexe avec l'*O*-acétylserine (thiol) lyase est insensible à la cystéine. Ce résultat nous permet d'infirmer que cet ADNc (L34076, Figure 2) code pour une serine acétyltransférase cytosolique, car la composition en acides aminés de l'extrémité N-terminale ne présente pas les caractéristiques de peptides de transit, et de plus cette sérine acétyltransférase est insensible à la cystéine. Ce dernier résultat parallèle les observations obtenues pour l'activité serine acétyltransférase cytosolique de feuilles de pois (Figure 1 et Tableau I).

**Exemple 4. Isolement et caractérisation d'un gène codant pour une isoforme serine acétyltransférase cytoplasmique (SAT3')**

On reprend le mode opératoire de l'exemple 3 avec les oligonucléotides 3 et 4 suivants :

15 Oligo 3 : 5' GAGAGAGGAT CCTCTTATCG CCGCGTTAAT ATGCCACCGG  
CCGGAGAACTC C 3'

Oligo 4 : 5' GAGCCTTACC AGTCTAATGT AGTATATTTC AACCTCGAGA  
GAGAG 3'

20 On isole un gène codant pour une acétyltransférase (U 30298) représentée sur la figure 5 (SEQ ID NO 2).

**Exemple 5. Isolement et caractérisation de gènes codant pour une isoforme serine acétyltransférase putative chloroplastique (SAT1)**

Le mode opératoire décrit pour l'exemple 3 est repris pour le présent exemple.

25 Un gène codant pour une serine acétyltransférase putative chloroplastique (L78443) représenté sur la figure 6 (SEQ ID NO 3) a été isolée par complémentation fonctionnelle d'une souche d'*Escherichia coli* déficiente pour l'activité serine acétyltransférase. [9] L'analyse de la séquence primaire a montré la présence de fortes similitudes avec la séquence de l'enzyme bactérienne (40% d'identité).

30 Les amorces qui ont été utilisées pour l'amplification de la séquence nucléotidique et son clonage dans le vecteur utilisé pour la transformation de plants de tabac (en caractères gras sur la figure 3) sont les suivantes :

Oligo 5 : 5' GAGAGAGGAT CCCCTCCTCC TCCTCCTCCT ATGGCTGCGT  
GCATCGACAC CTG 3'

35 Oligo 6 : 5' GCTCACCAGC CTAATACATT AAACCTTTTTC AGCTCGAGAG  
AGAG 2'

Ces amorces permettent l'introduction des sites de restriction *Bam*HI en 5' (GGATCC) et *Sac*I en 3' (GAGCTC).

On obtient un deuxième gène codant pour une serine acétyltransférase putative chloroplastique (U22964) représenté sur la figure 7 (SEQ ID NO 4) en reprenant le même mode opératoire avec l'oligo 7 en remplacement de l'oligo 5 comme amorce en 5'.

5 Oligo 7°: 5' GAGAGAGGAT CCGGCCGAGA AAAAAAAAAA ATGTTGCCGG  
TCACAAGTCG CCG 3'

#### Exemple 6. Constructions utilisées pour la transformation des plants de tabac variété petit Havanna

##### 10 Expression du transgène dans les feuilles

Les transformations des plants de tabac est réalisé par l'intermédiaire d'*Agrobacterium tumefaciens* EHA105, contenant le vecteur pBI121 (Clonetch) (Figures 4 et 5).

##### SAT3

15 Afin d'obtenir une expression de la SAT3 (SEQ ID NO 1) de l'exemple 2 dans le chloroplaste (figure 2), on introduira en position 5' de L'ADNc une extension qui permettra l'adressage dans ce compartiment. Pour cela, le peptide de transit optimisé décrit auparavant sera utilisé.

Entre les bordures gauche (BG) et droite (BD) du T-DNA sont clonés un gène de  
20 résistance à la kanamycine (NPTII) codant pour la néomycine phosphotransférase utilisés comme marqueur de sélection pour la transformation du tabac. L'expression du gene NPTII est sous la dépendance du promoteur et du terminateur de la nopaline synthase de *A. tumefaciens*. En aval, le gène de la glucuronidase cloné entre les sites uniques *Bam*H1 et *Sac*I, est sous le contrôle du promoteur 35S du virus de la mosaïque  
25 du chou-fleur (CaMV) et le signal de polyadénylation du gène de la nopaline synthase du plasmide Ti. L'insertion du produit de la OTP-SAT3 s'effectue entre les sites *Bam*H1 et *Sac*I du vecteur délété du gène de la  $\beta$ -glucuronidase.

##### SAT3'

On reprend le même mode opératoire que précédemment en remplaçant  
30 l'enzyme SAT3 (SEQ ID NO 1) de l'exemple 2 par l'enzyme SAT3' (SEQ ID NO 2) de l'exemple 4.

##### SAT1

Entre les bordures gauche (BG) et droite (BD) du T-DNA sont clonés un gène de  
résistance à la kanamycine (NPTII) codant pour la néomycine phosphotransférase  
35 utilisés comme marqueur de sélection pour la transformation du tabac. L'expression du gene NPTII est sous la dépendance du promoteur et du terminateur de la nopaline synthase de *A. tumefaciens*. En aval, le gene de la glucuronidase cloné entre les sites uniques *Bam*H1 et *Sac*I, est sous le contrôle du promoteur 35S du virus de la mosaïque

du chou-fleur (CaMV) et le signal de polyadénylation du gène de la nopaline synthase du plasmide Ti. L'insertion du gène codant pour une SAT1 (SEQ ID NO 3 ou 4) s'effectue entre les sites *Bam*HI et *Sac*I du vecteur délété du gène de la  $\beta$ -glucuronidase.

### Expression des transgènes dans les graines

Une construction similaire à celle présentée dans la figure 3 est réalisée dans le but d'obtenir une expression spécifique du transgène dans les graines. Cette stratégie pourrait être importante puisque les graines composent l'apport principale pour l'alimentation animale. Pour cela le promoteur constitutif de la mosaïque du tabac sera  
5 remplacé par un promoteur qui permet une expression spécifique du transgène pendant la mise en place des réserves de la graines.

#### **Exemple 6. Transformation du tabac**

10 De jeunes feuilles de plants de tabac (âgés de 3 à 4 semaines) dont la surface est stérilisée avec une solution de javel 10% (V/V) pendant 10 min puis rincée à l'eau stérile, sont découpées à l'emporte-pièce (30 disques par construction). 20 ml d'une culture de 48 heures d'*Agrobacterium tumefaciens* EHA105 (contenant le vecteur pBI121 modifié selon l'invention) sont centrifugés puis resuspendus dans 4 ml d'une  
15 solution de  $MgSO_4$  10mM. Les disques foliaires sont incubés pendant quelques minutes avec la solution d'agrobactéries puis transférés sur milieu MS (Sigma M-5519) supplémenté avec 0,05 mg/l d'acide  $\alpha$ -naphtalène acétique (NAA, Sigma), 2 mg/l de 6-benzylaminopurine (BAP) et 7 mg/l de phytoagar, pendant 2 à 3 jours. Les disques foliaires sont ensuite transférés sur un milieu identique auquel sont ajoutés 350 mg/l de  
20 cefotaxime (bacteriostatique) et 75 mg/l de kanamycine (agent de sélection). Au bout de 2 semaines, les disques sur lesquels se sont développés des cals ainsi que de jeunes pousses sont repiqués sur un milieu identique afin d'accélérer la croissance des pousses. Une semaine plus tard, les pousses vertes sont excisées et transférées sur le même milieu sans hormones, afin de permettre le développement des racines, ceci pendant 2  
25 semaines environ, au bout desquelles les jeunes plantes sont transférées en terre et cultivées en serre.

#### **Exemple 7. Analyse des résultats pour les plantes transgéniques et témoins.**

L'impact de l'expression de SAT3, SAT et SAT1 dans les plants de tabac est  
30 analysé au niveau du contenu en composés soufrés, la cystéine et la méthionine et le glutathion. La cystéine et le glutathion sont mis en évidence par la Méthode de Fahey ([23] Fahey, R.C. and Newton, G.L., Methods Enzymol. (1987) 143, 85-96), après dérivatisation des composés par le thiolyle (Calbiochem) et séparation par chromatographie liquide haute performance (CLHP) [23]. Le dosage de la méthionine  
35 libre est effectué par les méthodes de dosage des acides aminés libres après extraction et dérivation par l'orthophtaldéhyde et séparation par CLHP ([24] Brunet, P. & al., J. Chrom. (1988) 455, 173-182). L'activité de la serine acetyltransferase s'effectue comme décrit dans la méthodologie de dosage de l'O-acetylserine formée, par la technique

HPLC, ou par la méthode de couplage en présence d'un excès d'O-acétylserine (thiol) lyase [9], [14].

- 5 L'expression du gène de la serine acétyltransférase d'*Arabidopsis thaliana* dans le tabac conduit à une augmentation du taux en cystéine, du taux en glutathion et du taux en méthionine dans les tissus des plantes transformées par rapport aux plantes contrôles. Cette augmentation en composés soufrés s'accompagne d'une augmentation du contenu en d'autres acides aminés essentiels comme la threonine, isoleucine, leucine valine, tyrosine et phenylalanine. Toutes les augmentations en acides aminés sont corrélées avec une augmentation en l'activité serine acétyltransférase (SAT3, SAT3' ou
- 10 SAT1) liée à l'expression du transgène dans le chloroplaste. De plus, l'augmentation en ces composés soufrés conduit à améliorer le rapport nutritionnelle N/S de la plante et se traduit par un enrichissement des protéines de réserves de la graines en polypeptides riches en acides aminé soufrés au détriment des polypeptides pauvres en ces composés.

### Revendications

1. Procédé pour augmenter la production de cystéine par les cellules végétales et les plantes, ledit procédé consistant à surexprimer une SAT dans les chloroplastes des cellules végétales, et des plantes contenant les dites cellules végétales.
2. Procédé selon la revendication 1, caractérisé en ce que la SAT surexprimée dans les chloroplastes est une SAT sensible à la cystéine, comme par exemple une SAT chloroplastique de plante (SAT1) ou une SAT native d'origine bactérienne.
3. Procédé selon la revendication 1, caractérisé en ce que la SAT surexprimée dans les chloroplastes est une SAT insensible à la cystéine, comme par exemple une SAT cytoplasmique de plante (SAT3) ou une SAT d'origine bactérienne ou de plante mutée, rendue insensible à la cystéine par mutagenèse.
4. Procédé selon l'une des revendications 1 à 3, caractérisé en ce que la SAT est une SAT d'*Arabidopsis thaliana*.
5. Procédé selon l'une des revendications 1 à 4, caractérisé en ce que la protéine SAT est choisie parmi les protéines SAT3, SAT3' et SAT1 représentées avec leur séquence codante par les identificateurs de séquence 1, 2, et 3 et 4 respectivement (SEQ ID NO 1 à 4).
6. Procédé selon l'une des revendications 1 à 5, caractérisé en ce que la SAT est surexprimée dans les chloroplastes par intégration dans l'ADN chloroplastique des cellules végétales d'un gène chimère comprenant une séquence d'ADN codant pour ladite SAT, sous le contrôle d'éléments de régulations en 5' et 3' fonctionnels dans les chloroplastes.
7. Procédé selon l'une des revendications 1 à 5, caractérisé en ce que la SAT est surexprimée dans le cytoplasme sous la forme d'une protéine de fusion peptide de transit/SAT, la SAT mature fonctionnelle étant libérée à l'intérieur des chloroplastes.
8. Protéine de fusion peptide de transit/SAT, dans laquelle la SAT est définie dans l'une des revendications 2 à 5, et dans laquelle le peptide de transit est constitué par au moins un peptide de transit d'une protéine végétale naturelle à localisation plastidiale.
9. Protéine de fusion selon la revendication 8, caractérisée en ce que la SAT est homologue du peptide de transit.
10. Protéine de fusion selon la revendication 8, caractérisée en ce que la SAT est hétérologue du peptide de transit.
11. Protéine de fusion selon la revendication 10, caractérisée en ce que le peptide de transit est un peptide de transit d'une autre protéine à localisation plastidiale.



12. Protéine de fusion selon la revendication 11, caractérisé en ce que le peptide de transit est constitué par le peptide de transit d'une EPSPS de plante ou le peptide de transit d'une ssu RuBisCO de plante.

13. Protéine de fusion selon la revendication 10, caractérisé en ce que le peptide de transit comprend un peptide de transit d'une protéine végétale à localisation plastidiale et une partie de séquence de la partie mature N-terminale d'une protéine à localisation plastidiale entre la partie C-terminale du peptide de transit et la partie N-terminale de la SAT.

14. Protéine de fusion selon la revendication 13, caractérisée en ce que la partie de séquence comprend généralement moins de 40 acides aminés de la partie N-terminale de la protéine mature, de préférence moins de 30 acides aminés, plus préférentiellement entre 15 et 25 acides aminés.

15. Protéine de fusion selon l'une des revendications 13 ou 14, caractérisée en ce que le peptide de transit comprend entre la partie C-terminale de la partie N-terminale de la protéine mature et la partie N-terminale de la SAT un deuxième peptide de transit d'une protéine végétale à localisation plastidiale.

16. Protéine de fusion selon la revendication 15, caractérisée en ce que le peptide de transit est un peptide de transit optimisé (OTP) constitué par la fusion d'un premier peptide de transit, avec une partie de séquence de la partie mature N-terminale d'une protéine à localisation plastidiale, laquelle est fusionnée avec un deuxième peptide de transit.

17. Séquence d'acide nucléique codant pour une protéine de fusion peptide de transit/SAT selon l'une des revendications 8 à 16.

18. Gène chimère comprenant une séquence codante ainsi que des éléments de régulation en position 5' et 3' hétérologues pouvant fonctionner dans un organisme hôte, caractérisé en ce que la séquence codante comprend au moins une séquence d'acide nucléique selon la revendication 17.

19. Gène chimère selon la revendication 18, caractérisé en ce que l'organisme hôte est choisi parmi les bactéries, par exemple *E. coli*, les levures, en particulier des genres *Saccharomyces* ou *Kluyveromyces*, *Pichia*, les champignons, en particulier *Aspergillus*, les baculovirus, ou les cellules végétales et les plantes.

20. Gène chimère selon la revendication 19, caractérisé en ce que l'organisme hôte est une cellule végétale ou une plante la contenant.

21. Gène chimère selon la revendication 20, caractérisé en ce que l'élément de régulation en 5' comprend les séquences de régulation promotrice dans cellules végétales et les plantes, choisi parmi les promoteurs s'exprimant dans les feuilles des plantes, les promoteurs constitutifs, ou les promoteurs lumière dépendants d'origine bactérienne, virale ou végétale

22. Gène chimère selon la revendication 20, caractérisé en ce que l'élément de régulation en 5' comprend les séquences de régulation promotrice dans les cellules végétales et les plantes, choisi parmi les promoteurs spécifiques des graines.

23. Gène chimère selon la revendication 22, caractérisé en ce que le promoteur est choisi parmi les promoteurs de la napine, de la phaseoline, de la glutenine, de la zéine, de l'héliantinine, de l'albumine ou de l'oléosine.

24. Vecteur de clonage et/ou d'expression pour la transformation d'un organisme hôte caractérisé en ce qu'il contient au moins un gène chimère tel que défini selon l'une des revendications 18 à 23.

25. Procédé de transformation des organismes hôtes caractérisé en ce que l'on intègre dans le génome dudit organisme hôte au moins une séquence d'acide nucléique selon la revendication 17 ou un gène chimère selon l'une des revendications 18 à 23.

26. Procédé selon la revendication 25, au moyen du vecteur selon la revendication 24.

27. Procédé selon l'une des revendications 25 ou 26, caractérisé en ce que l'organisme hôte est choisi parmi les bactéries, par exemple *E. coli*, les levures, en particulier des genres *Saccharomyces* ou *Kluyveromyces*, *Pichia*, les champignons, en particulier *Aspergillus*, les baculovirus, ou les cellules végétales et les plantes.

28. Procédé selon la revendication 27, caractérisé en ce que l'organisme hôte est une cellule végétale ou une plante la contenant.

29. Procédé selon la revendication 28, caractérisé en ce que la plante est régénérée à partir d'une cellule végétale transformée.

30. Procédé selon la revendication 29, caractérisé en ce que l'organisme hôte est une plante monocotylédone, en particulier choisie parmi les céréales, la canne à sucre, le riz et le maïs, ou une plante dicotylédone, en particulier choisie parmi le tabac, la soja, le colza, le coton, la betterave et le trèfle.

31. Organisme hôte transformé, caractérisé en ce qu'il comprend au moins une séquence d'acide nucléique selon la revendication 17 ou un gène chimère selon l'une des revendications 18 à 23.

32. Organisme hôte selon la revendication 30, caractérisé en ce qu'il est obtenu par le procédé selon l'une des revendications 25 à 29.

32. Cellule végétale, caractérisée en ce qu'elle comprend au moins une séquence d'acide nucléique selon la revendication 17 ou un gène chimère selon l'une des revendications 18 à 23.

33. Plante génétiquement modifiée, caractérisée en ce qu'elle comprend au moins une cellule végétale selon la revendication 32.

34. Plante selon la revendication 33, caractérisé en ce que la plante est régénérée à partir d'une cellule végétale selon la revendication 32.

35. Plante génétiquement modifiée, caractérisée en ce qu'elle est issue de la culture et/ou du croisement des plantes régénérées selon la revendication 34.

36. Plante génétiquement modifiée selon l'une des revendications 33 à 35, caractérisée en ce qu'elle est une plante monocotylédone, en particulier choisie parmi les  
5 céréales, la canne à sucre, le riz et le maïs, ou une plante dicotylédone, en particulier choisie parmi le tabac, la soja, le colza, le coton, la betterave et le trèfle.

37. Plante génétiquement modifiée selon, l'une des revendications 33 à 36, caractérisée en ce qu'elle comprend d'autres gènes d'intérêt.

38. Plante selon la revendication 37, caractérisée en ce qu'elle comprend au  
10 moins un autre gène modifiant la teneur et la qualité des protéines de ladite plante, en particuliers dans les feuilles et/ou les graines.

39. Plante selon la revendication 38, caractérisée en ce que le gène code pour une protéine enrichie en acides aminés soufrés.

40. Graines des plantes génétiquement modifiées selon l'une des  
15 revendications 33 à 39.

# LISTE DE SEQUENCES

## (1) INFORMATIONS GENERALES:

### (i) DEPOSANT:

- (A) NOM: RHONE-POULENC AGROCHIMIE
- (B) RUE: 14-20 Rue Pierre BAISET
- (C) VILLE: LYON
- (E) PAYS: France
- (F) CODE POSTAL: 69009
- (G) TELEPHONE: 04 72 85 25 92
- (H) TELECOPIE: 04 72 85 28 43

(ii) TITRE DE L' INVENTION: Procédé pour augmenter la teneur en cystéine, méthionine et glutathion chez les plantes et plantes obtenues

(iii) NOMBRE DE SEQUENCES: 11

### (iv) FORME DECHIFFRABLE PAR ORDINATEUR:

- (A) TYPE DE SUPPORT: Floppy disk
- (B) ORDINATEUR: IBM PC compatible
- (C) SYSTEME D' EXPLOITATION: PC-DOS/MS-DOS
- (D) LOGICIEL: PatentIn Release #1.0, Version #1.30 (OEB)

## (2) INFORMATIONS POUR LA SEQ ID NO: 1:

### (i) CARACTERISTIQUES DE LA SEQUENCE:

- (A) LONGUEUR: 984 paires de bases
- (B) TYPE: nucléotide
- (C) NOMBRE DE BRINS: simple
- (D) CONFIGURATION: linéaire

(ii) TYPE DE MOLECULE: ADN

### (ix) CARACTERISTIQUE:

- (A) NOM/CLE: CDS
- (B) EMBLACEMENT: 31..972

### (xi) DESCRIPTION DE LA SEQUENCE: SEQ ID NO: 1:

GAGAGAGGAT CCTCTTTCCA ATCATAAACC ATG GCA ACA TGC ATA GAC ACA TGC	54
Met Ala Thr Cys Ile Asp Thr Cys	
1 5	
CGA ACC GGT AAT ACC CAA GAC GAT GAT TCC CGG TTC TGT TGC ATC AAG	102
Arg Thr Gly Asn Thr Gln Asp Asp Ser Arg Phe Cys Cys Ile Lys	
10 15 20	
AAT TTC TTT CGA CCC GGT TTC TCT GTA AAC CGG AAG ATT CAC CAC ACC	150
Asn Phe Phe Arg Pro Gly Phe Ser Val Asn Arg Lys Ile His His Thr	
25 30 35 40	
CAA ATC GAA GAT GAC GAT GAT GTC TGG ATC AAG ATG CTT GAA GAA GCC	198
Gln Ile Glu Asp Asp Asp Val Trp Ile Lys Met Leu Glu Glu Ala	
45 50 55	
AAA TCC GAT GTT AAA CAA GAA CCC ATT TTA TCA AAC TAC TAC TAC GCT	246
Lys Ser Asp Val Lys Gln Glu Pro Ile Leu Ser Asn Tyr Tyr Tyr Ala	
60 65 70	
TCG ATC ACA TCT CAT CGA TCT TTA GAG TCT GCT TTA GCT CAC ATC CTC	294
Ser Ile Thr Ser His Arg Ser Leu Glu Ser Ala Leu Ala His Ile Leu	
75 80 85	
TCC GTA AAG CTC AGC AAT TTA AAC CTA CCA AGC AAC ACA CTC TTC GAA	342
Ser Val Lys Leu Ser Asn Leu Asn Leu Pro Ser Asn Thr Leu Phe Glu	
90 95 100	

CTG TTC ATA AGC GTT TTA GAA GAA AGC CCT GAG ATC ATC GAA TCC ACG Leu Phe Ile Ser Val Leu Glu Glu Ser Pro Glu Ile Ile Glu Ser Thr 105 110 115 120	390
AAG CAA GAT CTT ATA GCA GTC AAA GAA AGA GAC CCA GCT TGT ATA AGC Lys Gln Asp Leu Ile Ala Val Lys Glu Arg Asp Pro Ala Cys Ile Ser 125 130 135	438
TAC GTT CAT TGC TTC TTG GGC TTC AAA GGC TTC CTC GCT TGT CAA GCT Tyr Val His Cys Phe Leu Gly Phe Lys Gly Phe Leu Ala Cys Gln Ala 140 145 150	486
CAT CGA ATA GCT CAT ACC CTC TGG AAA CAG AAC AGA AAA ATC GTA GCT His Arg Ile Ala His Thr Leu Trp Lys Gln Asn Arg Lys Ile Val Ala 155 160 165	534
TTA TTG ATC CAA AAC AGA GTA TCA GAA TCT TTC GCC GTC GAT ATT CAT Leu Leu Ile Gln Asn Arg Val Ser Glu Ser Phe Ala Val Asp Ile His 170 175 180	582
CCC GGA GCG AAG ATC GGA AAA GGG ATT CTT TTA GAC CAT GCG ACG GGC Pro Gly Ala Lys Ile Gly Lys Gly Ile Leu Leu Asp His Ala Thr Gly 185 190 195 200	630
GTG GTG ATC GGA GAG ACG GCG GTG GTT GGA GAC AAT GTT TCG ATT CTA Val Val Ile Gly Glu Thr Ala Val Val Gly Asp Asn Val Ser Ile Leu 205 210 215	678
CAC GGA GTG ACC TTG GGA GGA ACA GGG AAA CAG AGT GGT GAT CGG CAT His Gly Val Thr Leu Gly Gly Thr Gly Lys Gln Ser Gly Asp Arg His 220 225 230	726
CCG AAG ATT GGT GAT GGT GTG TTG ATT GGA GCT GGG AGT TGT ATA TTG Pro Lys Ile Gly Asp Gly Val Leu Ile Gly Ala Gly Ser Cys Ile Leu 235 240 245	774
GGG AAT ATA ACA ATC GGT GAG GGA GCT AAG ATT GGA TCA GGG TCG GTG Gly Asn Ile Thr Ile Gly Glu Gly Ala Lys Ile Gly Ser Gly Ser Val 250 255 260	822
GTG GTT AAG GAT GTG CCG GCG CGT ACG ACG GCG GTT GGA AAT CCG GCG Val Val Lys Asp Val Pro Ala Arg Thr Thr Ala Val Gly Asn Pro Ala 265 270 275 280	870
AGG TTG ATT GGT GGG AAA GAG AAT CCG AGA AAA CAT GAT AAG ATT CCT Arg Leu Ile Gly Gly Lys Glu Asn Pro Arg Lys His Asp Lys Ile Pro 285 290 295	918
TGT CTG ACT ATG GAC CAG ACA TCG TAT TTA ACC GAG TGG TCT GAT TAT Cys Leu Thr Met Asp Gln Thr Ser Tyr Leu Thr Glu Trp Ser Asp Tyr 300 305 310	966
GTG ATT TAACACAAAT GT Val Ile	984

(2) INFORMATIONS POUR LA SEQ ID NO: 2:

- (i) CARACTERISTIQUES DE LA SEQUENCE:
- (A) LONGUEUR: 974 paires de bases
  - (B) TYPE: nucléotide
  - (C) NOMBRE DE BRINS: simple
  - (D) CONFIGURATION: linéaire

(ii) TYPE DE MOLECULE: ADN

- (ix) CARACTERISTIQUE:
- (A) NOM/CLE: CDS
  - (B) EMBLACEMENT: 31..966

## (xi) DESCRIPTION DE LA SEQUENCE: SEQ ID NO: 2:

GAGAGAGGAT CCTCTTATCG CCGCGTTAAT ATG CCA CCG GCC GGA GAA CTC CGA	54
Met Pro Pro Ala Gly Glu Leu Arg	
1 5	
CAT CAA TCT CCA TCA AAG GAG AAA CTA TCT TCC GTT ACC CAA TCC GAT	102
His Gln Ser Pro Ser Lys Glu Lys Leu Ser Ser Val Thr Gln Ser Asp	
10 15 20	
GAA GCA GAA GCA GCG TCA GCA GCG ATA TCT GCG GCA GCT GCA GAT GCG	150
Glu Ala Glu Ala Ala Ser Ala Ala Ile Ser Ala Ala Ala Asp Ala	
25 30 35 40	
GAA GCT GCC GGA TTA TGG ACA CAG ATC AAG GCG GAA GCT CGC CGT GAT	198
Glu Ala Ala Gly Leu Trp Thr Gln Ile Lys Ala Glu Ala Arg Arg Asp	
45 50 55	
GCT GAG GCG GAG CCA GCT TTA GCT AGC TAT CTA TAT TCG ACG ATT CTT	246
Ala Glu Ala Glu Pro Ala Leu Ala Ser Tyr Leu Tyr Ser Thr Ile Leu	
60 65 70	
TCT CAT TCG TCT CTT GAA CGA TCT ATC TCG TTT CAT CTA GGA AAC AAG	294
Ser His Ser Ser Leu Glu Arg Ser Ile Ser Phe His Leu Gly Asn Lys	
75 80 85	
CTT TGT TCC TCA ACG CTT TTA TCC ACA CTT TTA TAC GAT CTG TTC TTA	342
Leu Cys Ser Ser Thr Leu Leu Ser Thr Leu Leu Tyr Asp Leu Phe Leu	
90 95 100	
AAC ACT TTT TCC TCC GAT CCT TCT CTT CGT AAC GCC ACC GTC GCA GAT	390
Asn Thr Phe Ser Ser Asp Pro Ser Leu Arg Asn Ala Thr Val Ala Asp	
105 110 115 120	
CTA CGC GCT GCT CGT GTT CGT GAT CCT GCT TGT ATC TCG TTC TCT CAT	438
Leu Arg Ala Ala Arg Val Arg Asp Pro Ala Cys Ile Ser Phe Ser His	
125 130 135	
TGT CTC CTC AAT TAC AAA GGC TTC TTA GCT ATT CAG GCG CAT CGT GTA	486
Cys Leu Leu Asn Tyr Lys Gly Phe Leu Ala Ile Gln Ala His Arg Val	
140 145 150	
TCA CAC AAG CTA TGG ACA CAA TCA CGG AAG CCA TTA GCA TTA GCT CTA	534
Ser His Lys Leu Trp Thr Gln Ser Arg Lys Pro Leu Ala Leu Ala Leu	
155 160 165	
CAC TCA AGA ATC TCC GAT GTA TTC GCT GTT GAT ATC CAT CCA GCA GCG	582
His Ser Arg Ile Ser Asp Val Phe Ala Val Asp Ile His Pro Ala Ala	
170 175 180	
AAG ATC GGA AAA GGG ATA CTT CTA GAC CAC GCA ACC GGA GTT GTA GTC	630
Lys Ile Gly Lys Gly Ile Leu Leu Asp His Ala Thr Gly Val Val Val	
185 190 195 200	
GGA GAA ACA GCG GTG ATT GGG AAC AAT GTT TCA ATC CTT CAC CAT GTG	678
Gly Glu Thr Ala Val Ile Gly Asn Asn Val Ser Ile Leu His His Val	
205 210 215	
ACA CTA GGT GGA ACA GGT AAA GCT TGT GGA GAT AGA CAT CCG AAG ATC	726
Thr Leu Gly Gly Thr Gly Lys Ala Cys Gly Asp Arg His Pro Lys Ile	
220 225 230	
GGT GAC GGT TGT TTG ATT GGA GCT GGA GCG ACT ATT CTT GGA AAT GTG	774
Gly Asp Gly Cys Leu Ile Gly Ala Gly Ala Thr Ile Leu Gly Asn Val	
235 240 245	
AAG ATT GGT GCA GGT GCT AAA GTA GGA GCT GGT TCT GTT GTG CTG ATT	822
Lys Ile Gly Ala Gly Ala Lys Val Gly Ala Gly Ser Val Val Leu Ile	
250 255 260	

GAC GTG CCT TGT CGA GGT ACT GCG GTT GGG AAT CCG GCG AGA CTT GTC	870
Asp Val Pro Cys Arg Gly Thr Ala Val Gly Asn Pro Ala Arg Leu Val	
265 270 275 280	
GGA GGG AAA GAG AAG CCA ACG ATT CAT GAT GAG GAA TGT CCT GGA GAA	918
Gly Gly Lys Glu Lys Pro Thr Ile His Asp Glu Glu Cys Pro Gly Glu	
285 290 295	
TCG ATG GAT CAT ACT TCA TTC ATC TCG GAA TGG TCA GAT TAC ATC ATA	966
Ser Met Asp His Thr Ser Phe Ile Ser Glu Trp Ser Asp Tyr Ile Ile	
300 305 310	
TAAAGTTG	974

(2) INFORMATIONS POUR LA SEQ ID NO: 3:

- (i) CARACTERISTIQUES DE LA SEQUENCE:  
 (A) LONGUEUR: 1048 paires de bases  
 (B) TYPE: nucléotide  
 (C) NOMBRE DE BRINS: simple  
 (D) CONFIGURATION: linéaire

(ii) TYPE DE MOLECULE: ADN

- (ix) CARACTERISTIQUE:  
 (A) NOM/CLE: CDS  
 (B) EMLACEMENT: 31..1038

(xi) DESCRIPTION DE LA SEQUENCE: SEQ ID NO: 3:

GAGAGAGGAT CCCCTCCTCC TCCTCCTCCT ATG GCT GCG TGC ATC GAC ACC TGC	54
Met Ala Ala Cys Ile Asp Thr Cys	
1 5	
CGC ACT GGT AAA CCC CAG ATT TCT CCT CGC GAT TCT TCT AAA CAC CAC	102
Arg Thr Gly Lys Pro Gln Ile Ser Pro Arg Asp Ser Ser Lys His His	
10 15 20	
GAC GAT GAA TCT GGC TTT CGT TAC ATG AAC TAC TTC CGT TAT CCT GAT	150
Asp Asp Glu Ser Gly Phe Arg Tyr Met Asn Tyr Phe Arg Tyr Pro Asp	
25 30 35 40	
CGA TCT TCC TTC AAT GGA ACC CAG ACC AAA ACC CTC CAT ACT CGT CCT	198
Arg Ser Ser Phe Asn Gly Thr Gln Thr Lys Thr Leu His Thr Arg Pro	
45 50 55	
TTG CTT GAA GAT CTC GAT CGC GAC GCT GAA GTC GAT GAT GTT TGG GCC	246
Leu Leu Glu Asp Leu Asp Arg Asp Ala Glu Val Asp Asp Val Trp Ala	
60 65 70	
AAA ATC CGA GAA GAG GCT AAA TCT GAT ATC GCC AAA GAA CCT ATT GTT	294
Lys Ile Arg Glu Glu Ala Lys Ser Asp Ile Ala Lys Glu Pro Ile Val	
75 80 85	
TCC GCT TAT TAT CAC GCT TCG ATT GTT TCT CAG CGT TCG TTG GAA GCT	342
Ser Ala Tyr Tyr His Ala Ser Ile Val Ser Gln Arg Ser Leu Glu Ala	
90 95 100	
GCG TTG GCG AAT ACT TTA TCT GTT AAA CTC AGC AAT TTG AAT CTT CCA	390
Ala Leu Ala Asn Thr Leu Ser Val Lys Leu Ser Asn Leu Asn Leu Pro	
105 110 115 120	
AGC AAC ACG CTT TTC GAT TTG TTC TCT GGT GTT CTT CAA GGA AAC CCA	438
Ser Asn Thr Leu Phe Asp Leu Phe Ser Gly Val Leu Gln Gly Asn Pro	
125 130 135	
GAT ATT GTT GAA TCT GTC AAG CTA GAT CTT TTA GCT GTT AAG GAG AGA	486
Asp Ile Val Glu Ser Val Lys Leu Asp Leu Leu Ala Val Lys Glu Arg	
140 145 150	

GAT CCT GCT TGT ATA AGC TAC GTT CAT TGT TTC CTT CAC TTT AAA GGC	534
Asp Pro Ala Cys Ile Ser Tyr Val His Cys Phe Leu His Phe Lys Gly	
155 160 165	
TTC CTC GCT TGT CAA GCG CAT CGT ATT GCT CAT GAG CTT TGG ACT CAG	582
Phe Leu Ala Cys Gln Ala His Arg Ile Ala His Glu Leu Trp Thr Gln	
170 175 180	
GAC AGA AAA ATC CTA GCT TTG TTG ATC CAG AAC AGA GTC TCT GAA GCC	630
Asp Arg Lys Ile Leu Ala Leu Leu Ile Gln Asn Arg Val Ser Glu Ala	
185 190 195 200	
TTC GCT GTT GAT TTC CAC CCT GGA GCT AAA ATC GGT ACC GGG ATT TTG	678
Phe Ala Val Asp Phe His Pro Gly Ala Lys Ile Gly Thr Gly Ile Leu	
205 210 215	
CTA GAC CAT GCT ACG GCT ATT GTG ATC GGT GAG ACG GCG GTT GTG GGG	726
Leu Asp His Ala Thr Ala Ile Val Ile Gly Glu Thr Ala Val Val Gly	
220 225 230	
AAC AAT GTT TCG ATT CTC CAT AAC GTT ACG CTT GGA GGA ACG GGG AAA	774
Asn Asn Val Ser Ile Leu His Asn Val Thr Leu Gly Gly Thr Gly Lys	
235 240 245	
CAG TGT GGA GAT AGG CAC CCG AAG ATT GGC GAT GGG GTT TTG ATT GGA	822
Gln Cys Gly Asp Arg His Pro Lys Ile Gly Asp Gly Val Leu Ile Gly	
250 255 260	
GCT GGG ACT TGT ATT TTG GGG AAT ATC ACG ATT GGT GAA GGA GCT AAG	870
Ala Gly Thr Cys Ile Leu Gly Asn Ile Thr Ile Gly Glu Gly Ala Lys	
265 270 275 280	
ATT GGT GCG GGG TCG GTG GTG TTG AAA GAC GTG CCG CCG CGT ACG ACG	918
Ile Gly Ala Gly Ser Val Val Leu Lys Asp Val Pro Pro Arg Thr Thr	
285 290 295	
GCT GTT GGA AAT CCG GCG AGG TTG CTT GGT GGT AAA GAT AAT CCG AAA	966
Ala Val Gly Asn Pro Ala Arg Leu Leu Gly Gly Lys Asp Asn Pro Lys	
300 305 310	
ACG CAT GAC AAG ATT CCT GGT TTG ACT ATG GAC CAG ACG TCG CAT ATA	1014
Thr His Asp Lys Ile Pro Gly Leu Thr Met Asp Gln Thr Ser His Ile	
315 320 325	
TCC GAG TGG TCG GAT TAT GTA ATT TGAAAAAGTC	1048
Ser Glu Trp Ser Asp Tyr Val Ile	
330 335	

(2) INFORMATIONS POUR LA SEQ ID NO: 4:

- (i) CARACTERISTIQUES DE LA SEQUENCE:
- (A) LONGUEUR: 1213 paires de bases
  - (B) TYPE: nucléotide
  - (C) NOMBRE DE BRINS: simple
  - (D) CONFIGURATION: linéaire

(ii) TYPE DE MOLECULE: ADN

- (ix) CARACTERISTIQUE:
- (A) NOM/CLE: CDS
  - (B) EMBLACEMENT: 31..1203

(xi) DESCRIPTION DE LA SEQUENCE: SEQ ID NO: 4:

GAGAGAGGAT CCGGCCGAGA AAAAAAAAAA ATG TTG CCG GTC ACA AGT CGC CGC	54
Met Leu Pro Val Thr Ser Arg Arg	
1 5	



CAC TTC ACA ATG TCC CTA TAT ATG CTC CGT TCA TCT TCT CCA CAC ATC His Phe Thr Met Ser Leu Tyr Met Leu Arg Ser Ser Pro His Ile 10 15 20	102
AAT CAT CAC TCT TTC CTT CTT CCT TCT TTT GTT TCC TCC AAA TTC AAA Asn His His Ser Phe Leu Leu Pro Ser Phe Val Ser Ser Lys Phe Lys 25 30 35 40	150
CAC CAT ACT TTA TCT CCT CCT CCT TCT CCT CCT CCT CCT CCT CCT ATG His His Thr Leu Ser Pro Pro Pro Ser Pro Pro Pro Pro Pro Pro Met 45 50 55	198
GCT GCG TGC ATC GAC ACC TGC CGC ACT GGT AAA CCC CAG ATT TCT CCT Ala Ala Cys Ile Asp Thr Cys Arg Thr Gly Lys Pro Gln Ile Ser Pro 60 65 70	246
CGC GAT TCT TCT AAA CAC CAC GAC GAT GAA TCT GGC TTT CGT TAC ATG Arg Asp Ser Ser Lys His His Asp Asp Glu Ser Gly Phe Arg Tyr Met 75 80 85	294
AAC TAC TTC CGT TAT CCT GAT CGA TCT TCC TTC AAT GGA ACC CAG ACC Asn Tyr Phe Arg Tyr Pro Asp Arg Ser Ser Phe Asn Gly Thr Gln Thr 90 95 100	342
AAA ACC CTC CAT ACT CGT CCT TTG CTT GAA GAT CTC GAT CGC GAC GCT Lys Thr Leu His Thr Arg Pro Leu Leu Glu Asp Leu Asp Arg Asp Ala 105 110 115 120	390
GAA GTC GAT GAT GTT TGG GCC AAA ATC CGA GAA GAG GCT AAA TCT GAT Glu Val Asp Asp Val Trp Ala Lys Ile Arg Glu Glu Ala Lys Ser Asp 125 130 135	438
ATC GCC AAA GAA CCT ATT GTT TCC GCT TAT TAT CAC GCT TCG ATT GTT Ile Ala Lys Glu Pro Ile Val Ser Ala Tyr Tyr His Ala Ser Ile Val 140 145 150	486
TCT CAG CGT TCG TTG GAA GCT GCG TTG GCG AAT ACT TTA TCT GTT AAA Ser Gln Arg Ser Leu Glu Ala Ala Leu Ala Asn Thr Leu Ser Val Lys 155 160 165	534
CTC AGC AAT TTG AAT CTT CCA AGC AAC ACG CTT TTC GAT TTG TTC TCT Leu Ser Asn Leu Asn Leu Pro Ser Asn Thr Leu Phe Asp Leu Phe Ser 170 175 180	582
GGT GTT CTT CAA GGA AAC CCA GAT ATT GTT GAA TCT GTC AAG CTA GAT Gly Val Leu Gln Gly Asn Pro Asp Ile Val Glu Ser Val Lys Leu Asp 185 190 195 200	630
CTT TTA GCT GTT AAG GAG AGA GAT CCT GCT TGT ATA AGC TAC GTT CAT Leu Leu Ala Val Lys Glu Arg Asp Pro Ala Cys Ile Ser Tyr Val His 205 210 215	678
TGT TTC CTT CAC TTT AAA GGC TTC CTC GCT TGT CAA GCG CAT CGT ATT Cys Phe Leu His Phe Lys Gly Phe Leu Ala Cys Gln Ala His Arg Ile 220 225 230	726
GCT CAT GAG CTT TGG ACT CAG GAC AGA AAA ATC CTA GCT TTG TTG ATC Ala His Glu Leu Trp Thr Gln Asp Arg Lys Ile Leu Ala Leu Leu Ile 235 240 245	774
CAG AAC AGA GTC TCT GAA GCC TTC GCT GTT GAT TTC CAC CCT GGA GCT Gln Asn Arg Val Ser Glu Ala Phe Ala Val Asp Phe His Pro Gly Ala 250 255 260	822
AAA ATC GGT ACC GGG ATT TTG CTA GAC CAT GCT ACG GCT ATT GTG ATC Lys Ile Gly Thr Gly Ile Leu Leu Asp His Ala Thr Ala Ile Val Ile 265 270 275 280	870

GGT GAG ACG GCG GTT GTG GGG AAC AAT GTT TCG ATT CTC CAT AAC GTT Gly Glu Thr Ala Val Val Gly Asn Asn Val Ser Ile Leu His Asn Val 285 290 295	918
ACG CTT GGA GGA ACG GGG AAA CAG TGT GGA GAT AGG CAC CCG AAG ATT Thr Leu Gly Gly Thr Gly Lys Gln Cys Gly Asp Arg His Pro Lys Ile 300 305 310	966
GGC GAT GGG GTT TTG ATT GGA GCT GGG ACT TGT ATT TTG GGG AAT ATC Gly Asp Gly Val Leu Ile Gly Ala Gly Thr Cys Ile Leu Gly Asn Ile 315 320 325	1014
ACG ATT GGT GAA GGA GCT AAG ATT GGT GCG GGG TCG GTG GTG TTG AAA Thr Ile Gly Glu Gly Ala Lys Ile Gly Ala Gly Ser Val Val Leu Lys 330 335 340	1062
GAC GTG CCG CCG CGT ACG ACG GCT GTT GGA AAT CCG GCG AGG TTG CTT Asp Val Pro Pro Arg Thr Thr Ala Val Gly Asn Pro Ala Arg Leu Leu 345 350 355 360	1110
GGT GGT AAA GAT AAT CCG AAA ACG CAT GAC AAG ATT CCT GGT TTG ACT Gly Gly Lys Asp Asn Pro Lys Thr His Asp Lys Ile Pro Gly Leu Thr 365 370 375	1158
ATG GAC CAG ACG TCG CAT ATA TCC GAG TGG TCG GAT TAT GTA ATT Met Asp Gln Thr Ser His Ile Ser Glu Trp Ser Asp Tyr Val Ile 380 385 390	1203
TGAAAAAGTC	1213

(2) INFORMATIONS POUR LA SEQ ID NO: 5:

- (i) CARACTERISTIQUES DE LA SEQUENCE:
  - (A) LONGUEUR: 54 paires de bases
  - (B) TYPE: nucléotide
  - (C) NOMBRE DE BRINS: simple
  - (D) CONFIGURATION: linéaire

(ii) TYPE DE MOLECULE: ADN

- (ix) CARACTERISTIQUE:
  - (A) NOM/CLE: oligonucléotide 1

(xi) DESCRIPTION DE LA SEQUENCE: SEQ ID NO: 5:

GAGAGAGGAT CCTCTTTCCA ATCATAAACC ATGGCAACAT GCATAGACAC ATGC

54

(2) INFORMATIONS POUR LA SEQ ID NO: 6:

- (i) CARACTERISTIQUES DE LA SEQUENCE:
  - (A) LONGUEUR: 46 paires de bases
  - (B) TYPE: nucléotide
  - (C) NOMBRE DE BRINS: simple
  - (D) CONFIGURATION: linéaire

(ii) TYPE DE MOLECULE: ADN

- (ix) CARACTERISTIQUE:
  - (A) NOM/CLE: oligonucléotide 2

(xi) DESCRIPTION DE LA SEQUENCE: SEQ ID NO: 6:

GGCTCACCAG ACTAATACAC TAAATTGTGT TTACCTCGAG AGAGAG 46

(2) INFORMATIONS POUR LA SEQ ID NO: 7:

- (i) CARACTERISTIQUES DE LA SEQUENCE:

- (A) LONGUEUR: 51 paires de bases
- (B) TYPE: nucléotide
- (C) NOMBRE DE BRINS: simple
- (D) CONFIGURATION: linéaire

(ii) TYPE DE MOLECULE: ADN

(ix) CARACTERISTIQUE:

- (A) NOM/CLE: oligonucléotide 3

(xi) DESCRIPTION DE LA SEQUENCE: SEQ ID NO: 7:

GAGAGAGGAT CCTCTTATCG CCGCGTTAAT ATGCCACCGG CCGGAGAACTC C

51

(2) INFORMATIONS POUR LA SEQ ID NO: 8:

(i) CARACTERISTIQUES DE LA SEQUENCE:

- (A) LONGUEUR: 45 paires de bases
- (B) TYPE: nucléotide
- (C) NOMBRE DE BRINS: simple
- (D) CONFIGURATION: linéaire

(ii) TYPE DE MOLECULE: ADN

(ix) CARACTERISTIQUE:

- (A) NOM/CLE: oligonucléotide 4

(xi) DESCRIPTION DE LA SEQUENCE: SEQ ID NO: 8:

GAGCCTTACC AGTCTAATGT AGTATATTTT AACCTCGAGA GAGAG

45

(2) INFORMATIONS POUR LA SEQ ID NO: 9:

(i) CARACTERISTIQUES DE LA SEQUENCE:

- (A) LONGUEUR: 53 paires de bases
- (B) TYPE: nucléotide
- (C) NOMBRE DE BRINS: simple
- (D) CONFIGURATION: linéaire

(ii) TYPE DE MOLECULE: ADN

(ix) CARACTERISTIQUE:

- (A) NOM/CLE: oligonucléotide 5

(xi) DESCRIPTION DE LA SEQUENCE: SEQ ID NO: 9:

GAGAGAGGAT CCCCTCCTCC TCCTCCTCCT ATGGCTGCGT GCATCGACAC CTG

53

(2) INFORMATIONS POUR LA SEQ ID NO: 10:

(i) CARACTERISTIQUES DE LA SEQUENCE:

- (A) LONGUEUR: 44 paires de bases
- (B) TYPE: nucléotide
- (C) NOMBRE DE BRINS: simple
- (D) CONFIGURATION: linéaire

(ii) TYPE DE MOLECULE: ADN

(ix) CARACTERISTIQUE:

- (A) NOM/CLE: oligonucléotide 6

(xi) DESCRIPTION DE LA SEQUENCE: SEQ ID NO: 10:

GCTCACCAGC CTAATACATT AAACTTTTTC AGCTCGAGAG AGAG

44

(2) INFORMATIONS POUR LA SEQ ID NO: 11:

- (i) CARACTERISTIQUES DE LA SEQUENCE:  
 (A) LONGUEUR: 53 paires de bases  
 (B) TYPE: nucléotide  
 (C) NOMBRE DE BRINS: simple  
 (D) CONFIGURATION: linéaire

(ii) TYPE DE MOLECULE: ADN

- (ix) CARACTERISTIQUE:  
 (A) NOM/CLE: oligonucléotide 7

(xi) DESCRIPTION DE LA SEQUENCE: SEQ ID NO: 11:

GAGAGAGGAT CCGGCCGAGA AAAAAAAAAA ATGTTGCCGG TCACAAGTCG CCG

53

ou de préférence des cellules végétales et des plantes.

Par "cellule végétale", on entend selon l'invention toute cellule issue d'une plante et pouvant constituer des tissus indifférenciés tels que des cals, des tissus différenciés tels que des embryons, des parties de plantes, des plantes ou des semences.

5 On entend par "plante" selon l'invention, tout organisme multicellulaire différencié capable de photosynthèse, en particulier monocotylédones ou dicotylédones, plus particulièrement des plantes de culture destinées ou non à l'alimentation animale ou humaine, comme le maïs, le blé, le colza, le soja, le riz, la canne à sucre, la betterave, le tabac, le coton, etc.

10 Les éléments de régulation nécessaires à l'expression de la séquence d'acide nucléique codant pour une protéine de fusion selon l'invention sont bien connus de l'homme du métier en fonction de l'organisme hôte. Ils comprennent notamment des séquences promotrices, des activateurs de transcription, des séquences terminatrices, y compris des codons start et stop. Les moyens et méthodes pour identifier et sélectionner  
15 les éléments de régulation sont bien connus de l'homme du métier et largement décrits dans la littérature.

L'invention concerne plus particulièrement la transformation des plantes. Comme séquence de régulation promotrice dans les plantes, on peut utiliser toute séquence promotrice d'un gène s'exprimant naturellement dans les plantes en particulier  
20 un promoteur s'exprimant notamment dans les feuilles des plantes, comme par exemple des promoteurs dits constitutifs d'origine bactérienne, virale ou végétale ou encore des promoteurs dits lumière dépendants comme celui d'un gène de la petite sous-unité de ribulose- biscarboxylase/oxygénase (RuBisCO) de plante ou tout promoteur convenable connu pouvant être utilisé. Parmi les promoteurs d'origine végétale on citera les  
25 promoteurs d'histone tels que décrits dans la demande EP 0 507 698, ou le promoteur d'actine de riz (US 5,641,876). Parmi les promoteurs d'un gène de virus de plante, on citera celui de la mosaïque du chou fleur (CAMV 19S ou 35S), ou le promoteur du circovirus (AU 689 311).

On peut encore utiliser une séquence de régulation promotrice spécifique de  
30 régions ou de tissus particuliers des plantes, et plus particulièrement des promoteurs spécifiques des graines ([22] Datla, R.& al., Biotechnology Ann. Rev. (1997) 3, 269-296), notamment les promoteurs de la napine (EP 255 378), de la phaseoline, de la glutenine, de la zéine, de l'héliantinine (WO 92/17580), de l'albumine (WO 98/45460), de l'oélosine (WO 98/45461), de l'ATS1 ou de l'ATS3 (PCT/US98/06978, déposée le  
35 20 octobre 1998, incorporée ici par référence).

Selon l'invention, on peut également utiliser, en association avec la séquence de régulation promotrice, d'autres séquences de régulation, qui sont situées entre le promoteur et la séquence codante, telles que des activateurs de transcription ("enhancer"), comme par exemple l'activateur de translation du virus de la mosaïque du

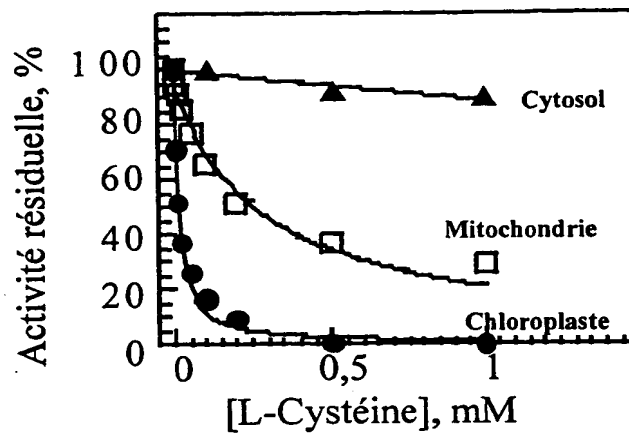


Figure 1 : Effet de la cystéine sur les activités sérine acétyltransférase de pois (*Pisum sativum*).

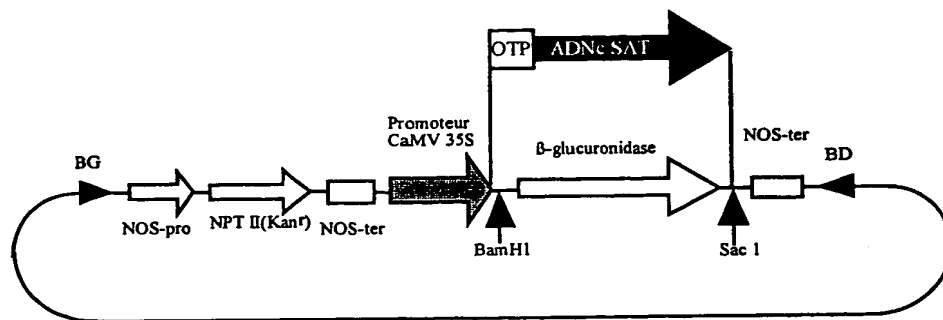


Figure 2: Procédure de clonage de l'OTP/Sérine acétyltransférase SAT3 ou SAT dans le vecteur pBI121.

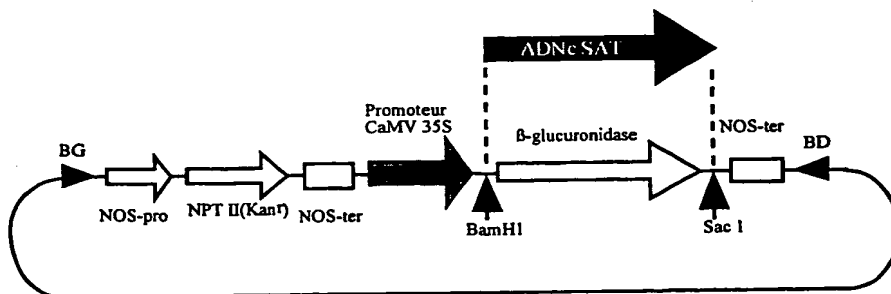


Figure 3: Procédure de clonage de la Sérine acétyltransférase SAT1 dans le vecteur pBI121.

																Bam HI	
																gagagaggatcctctttccaatcataaacc	
M	A	T	C	I	D	T	C	R	T	G	N	T	Q	D	D	16	
ATG	GCA	ACA	TGC	ATA	GAC	ACA	TGC	CGA	ACC	GGT	AAT	ACC	CAA	GAC	GAT	48	
D	S	R	F	C	C	I	K	N	F	F	R	P	G	F	S	32	
GAT	TCC	CGG	TTC	TGT	TGC	ATC	AAG	AAT	TTC	TTT	CGA	CCC	GGT	TTC	TCT	96	
V	N	R	K	I	H	H	T	Q	I	E	D	D	D	D	V	48	
GTA	AAC	CGG	AAG	ATT	CAC	CAC	ACC	CAA	ATC	GAA	GAT	GAC	GAT	GAT	GTC	144	
W	I	K	M	L	E	E	A	K	S	D	V	K	Q	E	P	64	
TGG	ATC	AAG	ATG	CTT	GAA	GAA	GCC	AAA	TCC	GAT	GTT	AAA	CAA	GAA	CCC	192	
I	L	S	N	Y	Y	Y	A	S	I	T	S	H	R	S	L	80	
ATT	TTA	TCA	AAC	TAC	TAC	TAC	GCT	TCG	ATC	ACA	TCT	CAT	CGA	TCT	TTA	240	
E	S	A	L	A	H	I	L	S	V	K	L	S	N	L	N	96	
GAG	TCT	GCT	TTA	GCT	CAC	ATC	CTC	TCC	GTA	AAG	CTC	AGC	AAT	TTA	AAC	288	
L	P	S	N	T	L	F	E	L	F	I	S	V	L	E	E	112	
CTA	CCA	AGC	AAC	ACA	CTC	TTC	GAA	CTG	TTC	ATA	AGC	GTT	TTA	GAA	GAA	336	
S	P	E	I	I	E	S	T	K	Q	D	L	I	A	V	K	128	
AGC	CCT	GAG	ATC	ATC	GAA	TCC	ACG	AAG	CAA	GAT	CTT	ATA	GCA	GTC	AAA	384	
E	R	D	P	A	C	I	S	Y	V	H	C	F	L	G	F	144	
GAA	AGA	GAC	CCA	GCT	TGT	ATA	AGC	TAC	GTT	CAT	TGC	TTC	TTG	GGC	TTC	432	
K	G	F	L	A	C	Q	A	H	R	I	A	H	T	L	W	160	
AAA	GGC	TTC	CTC	GCT	TGT	CAA	GCT	CAT	CGA	ATA	GCT	CAT	ACC	CTC	TGG	480	
K	Q	N	R	K	I	V	A	L	L	I	Q	N	R	V	S	176	
AAA	CAG	AAC	AGA	AAA	ATC	GTA	GCT	TTA	TTG	ATC	CAA	AAC	AGA	GTA	TCA	528	
E	S	F	A	V	D	I	H	P	G	A	K	I	G	K	G	192	
GAA	TCT	TTC	GCC	GTC	GAT	ATT	CAT	CCC	GGA	GCG	AAG	ATC	GGA	AAA	GGG	576	
I	L	L	D	H	A	T	G	V	V	I	G	E	T	A	V	208	
ATT	CTT	TTA	GAC	CAT	GCG	ACG	GGC	GTG	GTG	ATC	GGA	GAG	ACG	GCG	GTG	624	
V	G	D	N	V	S	I	L	H	G	V	T	L	G	G	T	224	
GTT	GGA	GAC	AAT	GTT	TCG	ATT	CTA	CAC	GGA	GTG	ACC	TTG	GGA	GGA	ACA	672	
G	K	Q	S	G	D	R	H	P	K	I	G	D	G	V	L	240	
GGG	AAA	CAG	AGT	GGT	GAT	CGG	CAT	CCG	AAG	ATT	GGT	GAT	GGT	GTG	TTG	720	
I	G	A	G	S	C	I	L	G	N	I	T	I	G	E	G	256	
ATT	GGA	GCT	GGG	AGT	TGT	ATA	TTG	GGG	AAT	ATA	ACA	ATC	GGT	GAG	GGA	768	
A	K	I	G	S	G	S	V	V	V	K	D	V	P	A	R	272	
GCT	AAG	ATT	GGA	TCA	GGG	TCG	GTG	GTG	GTT	AAG	GAT	GTG	CCG	GCG	CGT	816	
T	T	A	V	G	N	P	A	R	L	I	G	G	K	E	N	288	
ACG	ACG	GCG	GTT	GGA	AAT	CCG	GCG	AGG	TTG	ATT	GGT	GGG	AAA	GAG	AAT	864	
P	R	K	H	D	K	I	P	C	L	T	M	D	Q	T	S	304	
CCG	AGA	AAA	CAT	GAT	AAG	ATT	CCT	TGT	CTG	ACT	ATG	GAC	CAG	ACA	TCG	912	
Y	L	T	E	W	S	D	Y	V	I							314	
TAT	TTA	ACC	GAG	TGG	TCT	GAT	TAT	GTG	ATT	TAA	cac	aaa	tgt			945	
agagag																gg ctc acc aga cta ata cac taa att gtg ttt aca ctc gag	
																Sac I	

Figure 4: Séquence nucléotidique et peptidique du gène de l'isoforme SAT 3 (L34076)  
d'*A. thaliana*

BamHI															14
gagagaggatcctcttattcgccg															
cgttaat	M	P	P	A	G	E	L	R	H	Q	S	P	S	K	42
E K L S S V T Q S D E A E A A S	ATG	CCA	CCG	GCC	GGA	GAA	CTC	CGA	CAT	CAA	TCT	CCA	TCA	AAG	30
GAG AAA CTA TCT TCC GTT ACC CAA TCC GAT GAA GCA GAA GCA GCG TCA	ATA	TCT	TCC	GTT	ACC	CAA	TCC	GAT	GAA	GCA	GAA	GCA	GCG	TCA	90
A A I S A A A D A E A A G L W	ATA	TCT	GCG	GCA	GCT	GCA	GAT	GCG	GAA	GCT	GCC	GGA	TTA	TGG	46
GCA GCG ATA TCT GCG GCA GCT GCA GAT GCG GAA GCT GCC GGA TTA TGG	ATA	TCT	GCG	GCA	GCT	GCA	GAT	GCG	GAA	GCT	GCC	GGA	TTA	TGG	138
T Q I K A E A R R D A E A E P A	ATA	TCT	GCG	GCA	GCT	GCA	GAT	GCG	GAA	GCT	GCC	GGA	TTA	TGG	62
ACA CAG ATC AAG GCG GAA GCT CGC CGT GAT GCT GAG GCG GAG CCA GCT	ATA	TCT	GCG	GCA	GCT	GCA	GAT	GCG	GAA	GCT	GCC	GGA	TTA	TGG	186
L A S Y L Y S T I L S H S S L E	ATA	TCT	GCG	GCA	GCT	GCA	GAT	GCG	GAA	GCT	GCC	GGA	TTA	TGG	78
TTA GCT AGC TAT CTA TAT TCG ACG ATT CTT TCT CAT TCG TCT CTT GAA	ATA	TCT	GCG	GCA	GCT	GCA	GAT	GCG	GAA	GCT	GCC	GGA	TTA	TGG	234
R S I S F H L G N K L C S S T L	ATA	TCT	GCG	GCA	GCT	GCA	GAT	GCG	GAA	GCT	GCC	GGA	TTA	TGG	94
CGA TCT ATC TCG TTT CAT CTA GGA AAC AAG CTT TGT TCC TCA ACG CTT	ATA	TCT	GCG	GCA	GCT	GCA	GAT	GCG	GAA	GCT	GCC	GGA	TTA	TGG	282
L S T L L Y D L F L N T F S S D	ATA	TCT	GCG	GCA	GCT	GCA	GAT	GCG	GAA	GCT	GCC	GGA	TTA	TGG	110
TTA TCC ACA CTT TTA TAC GAT CTG TTC TTA AAC ACT TTT TCC TCC GAT	ATA	TCT	GCG	GCA	GCT	GCA	GAT	GCG	GAA	GCT	GCC	GGA	TTA	TGG	330
P S L R N A T V A D L R A A R V	ATA	TCT	GCG	GCA	GCT	GCA	GAT	GCG	GAA	GCT	GCC	GGA	TTA	TGG	126
CCT TCT CTT CGT AAC GCC ACC GTC GCA GAT CTA CGC GCT GCT CGT GTT	ATA	TCT	GCG	GCA	GCT	GCA	GAT	GCG	GAA	GCT	GCC	GGA	TTA	TGG	378
R D P A C I S F S H C L L N Y K	ATA	TCT	GCG	GCA	GCT	GCA	GAT	GCG	GAA	GCT	GCC	GGA	TTA	TGG	142
CGT GAT CCT GCT TGT ATC TCG TTC TCT CAT TGT CTC CTC AAT TAC AAA	ATA	TCT	GCG	GCA	GCT	GCA	GAT	GCG	GAA	GCT	GCC	GGA	TTA	TGG	426
G F L A I Q A H R V S H K L W T	ATA	TCT	GCG	GCA	GCT	GCA	GAT	GCG	GAA	GCT	GCC	GGA	TTA	TGG	158
GGC TTC TTA GCT ATT CAG GCG CAT CGT GTA TCA CAC AAG CTA TGG ACA	ATA	TCT	GCG	GCA	GCT	GCA	GAT	GCG	GAA	GCT	GCC	GGA	TTA	TGG	474
Q S R K P L A L A L H S R I S D	ATA	TCT	GCG	GCA	GCT	GCA	GAT	GCG	GAA	GCT	GCC	GGA	TTA	TGG	174
CAA TCA CGG AAG CCA TTA GCA TTA GCT CTA CAC TCA AGA ATC TCC GAT	ATA	TCT	GCG	GCA	GCT	GCA	GAT	GCG	GAA	GCT	GCC	GGA	TTA	TGG	522
V F A V D I H P A A K I G K G I	ATA	TCT	GCG	GCA	GCT	GCA	GAT	GCG	GAA	GCT	GCC	GGA	TTA	TGG	190
GTA TTC GCT GTT GAT ATC CAT CCA GCA GCG AAG ATC GGA AAA GGG ATA	ATA	TCT	GCG	GCA	GCT	GCA	GAT	GCG	GAA	GCT	GCC	GGA	TTA	TGG	570
L L D H A T G V V V G E T A V I	ATA	TCT	GCG	GCA	GCT	GCA	GAT	GCG	GAA	GCT	GCC	GGA	TTA	TGG	206
CTT CTA GAC CAC GCA ACC GGA GTT GTA GTC GGA GAA ACA GCG GTG ATT	ATA	TCT	GCG	GCA	GCT	GCA	GAT	GCG	GAA	GCT	GCC	GGA	TTA	TGG	618
G N N V S I L H H V T L G G T G	ATA	TCT	GCG	GCA	GCT	GCA	GAT	GCG	GAA	GCT	GCC	GGA	TTA	TGG	222
GGG AAC AAT GTT TCA ATC CTT CAC CAT GTG ACA CTA GGT GGA ACA GGT	ATA	TCT	GCG	GCA	GCT	GCA	GAT	GCG	GAA	GCT	GCC	GGA	TTA	TGG	666
K A C G D R H P K I G D G C L I	ATA	TCT	GCG	GCA	GCT	GCA	GAT	GCG	GAA	GCT	GCC	GGA	TTA	TGG	238
AAA GCT TGT GGA GAT AGA CAT CCG AAG ATC GGT GAC GGT TGT TTG ATT	ATA	TCT	GCG	GCA	GCT	GCA	GAT	GCG	GAA	GCT	GCC	GGA	TTA	TGG	714
G A G A T I L G N V K I G A G A	ATA	TCT	GCG	GCA	GCT	GCA	GAT	GCG	GAA	GCT	GCC	GGA	TTA	TGG	254
GGA GCT GGA GCG ACT ATT CTT GGA AAT GTG AAG ATT GGT GCA GGT GCT	ATA	TCT	GCG	GCA	GCT	GCA	GAT	GCG	GAA	GCT	GCC	GGA	TTA	TGG	762
K V G A G S V V L I D V P C R G	ATA	TCT	GCG	GCA	GCT	GCA	GAT	GCG	GAA	GCT	GCC	GGA	TTA	TGG	270
AAA GTA GGA GCT GGT TCT GTT GTG CTG ATT GAC GTG CCT TGT CGA GGT	ATA	TCT	GCG	GCA	GCT	GCA	GAT	GCG	GAA	GCT	GCC	GGA	TTA	TGG	810
T A V G N P A R L V G G K E K P	ATA	TCT	GCG	GCA	GCT	GCA	GAT	GCG	GAA	GCT	GCC	GGA	TTA	TGG	286
ACT GCG GTT GGG AAT CCG GCG AGA CTT GTC GGA GGG AAA GAG AAG CCA	ATA	TCT	GCG	GCA	GCT	GCA	GAT	GCG	GAA	GCT	GCC	GGA	TTA	TGG	858
T I H D E E C P G E S M D H T S	ATA	TCT	GCG	GCA	GCT	GCA	GAT	GCG	GAA	GCT	GCC	GGA	TTA	TGG	302
ACG ATT CAT GAT GAG GAA TGT CCT GGA GAA TCG ATG GAT CAT ACT TCA	ATA	TCT	GCG	GCA	GCT	GCA	GAT	GCG	GAA	GCT	GCC	GGA	TTA	TGG	906
F I S E W S D Y I I...	ATA	TCT	GCG	GCA	GCT	GCA	GAT	GCG	GAA	GCT	GCC	GGA	TTA	TGG	312
TTC ATC TCG GAA TGG TCA GAT TAC ATC ATA TAA AGT TG	ATA	TCT	GCG	GCA	GCT	GCA	GAT	GCG	GAA	GCT	GCC	GGA	TTA	TGG	939
g agc ctt acc agt cta atg tag tat att tca acc tca aga gag															
Sac 1															

ag

Figure 5: Séquence nucléotidique et peptidique du gène de l'isoforme SAT3' (U30298)  
d'*A. thaliana*



gagagaggatccccctcctcctcctcctcct																	
M	A	A	C	I	D	T	C	R	T	G	K	P	Q	I			15
ATG	GCT	GCG	TGC	ATC	GAC	ACC	TGC	CGC	ACT	GGT	AAA	CCC	CAG	ATT			45
S	P	R	D	S	S	K	H	H	D	D	E	S	G	F			30
TCT	CCT	CGC	GAT	TCT	TCT	AAA	CAC	CAC	GAC	GAT	GAA	TCT	GGC	TTT			90
R	Y	M	N	Y	F	R	Y	P	D	R	S	S	F	N			45
CGT	TAC	ATG	AAC	TAC	TTC	CGT	TAT	CCT	GAT	CGA	TCT	TCC	TTC	AAT			135
G	T	Q	T	K	T	L	H	T	R	P	L	L	E	D			60
GGA	ACC	CAG	ACC	AAA	ACC	CTC	CAT	ACT	CGT	CCT	TTG	CTT	GAA	GAT			180
L	D	R	D	A	E	V	D	D	V	W	A	K	I	R			75
CTC	GAT	CGC	GAC	GCT	GAA	GTC	GAT	GAT	GTT	TGG	GCC	AAA	ATC	CGA			225
E	E	A	K	S	D	I	A	K	E	P	I	V	S	A			90
GAA	GAG	GCT	AAA	TCT	GAT	ATC	GCC	AAA	GAA	CCT	ATT	GTT	TCC	GCT			270
Y	Y	H	A	S	I	V	S	Q	R	S	L	E	A	A			105
TAT	TAT	CAC	GCT	TCG	ATT	GTT	TCT	CAG	CGT	TCG	TTG	GAA	GCT	GCG			315
L	A	N	T	L	S	V	K	L	S	N	L	N	L	P			120
TTG	GCG	AAT	ACT	TTA	TCT	GTT	AAA	CTC	AGC	AAT	TTG	AAT	CTT	CCA			360
S	N	T	L	F	D	L	F	S	G	V	L	Q	G	N			135
AGC	AAC	ACG	CTT	TTC	GAT	TTG	TTC	TCT	GGT	GTT	CTT	CAA	GGA	AAC			405
P	D	I	V	E	S	V	K	L	D	L	L	A	V	K			150
CCA	GAT	ATT	GTT	GAA	TCT	GTC	AAG	CTA	GAT	CTT	TTA	GCT	GTT	AAG			450
E	R	D	P	A	C	I	S	Y	V	H	C	F	L	H			165
GAG	AGA	GAT	CCT	GCT	TGT	ATA	AGC	TAC	GTT	CAT	TGT	TTC	CTT	CAC			495
F	K	G	F	L	A	C	Q	A	H	R	I	A	H	E			180
TTT	AAA	GGC	TTC	CTC	GCT	TGT	CAA	GCG	CAT	CGT	ATT	GCT	CAT	GAG			540
L	W	T	Q	D	R	K	I	L	A	L	L	I	Q	N			195
CTT	TGG	ACT	CAG	GAC	AGA	AAA	ATC	CTA	GCT	TTG	TTG	ATC	CAG	AAC			585
R	V	S	E	A	F	A	V	D	F	H	P	G	A	K			210
AGA	GTC	TCT	GAA	GCC	TTC	GCT	GTT	GAT	TTC	CAC	CCT	GGA	GCT	AAA			630
I	G	T	G	I	L	L	D	H	A	T	A	I	V	I			225
ATC	GGT	ACC	GGG	ATT	TTG	CTA	GAC	CAT	GCT	ACG	GCT	ATT	GTG	ATC			675
G	E	T	A	V	V	G	N	N	V	S	I	L	H	N			240
GGT	GAG	ACG	GCG	GTT	GTG	GGG	AAC	AAT	GTT	TCG	ATT	CTC	CAT	AAC			720
V	T	L	G	G	T	G	K	Q	C	G	D	R	H	P			255
GTT	ACG	CTT	GGA	GGA	ACG	GGG	AAA	CAG	TGT	GGA	GAT	AGG	CAC	CCG			765
K	I	G	D	G	V	L	I	G	A	G	T	C	I	L			270
AAG	ATT	GGC	GAT	GGG	GTT	TTG	ATT	GGA	GCT	GGG	ACT	TGT	ATT	TTG			810
G	N	I	T	I	G	E	G	A	K	I	G	A	G	S			285
GGG	AAT	ATC	ACG	ATT	GGT	GAA	GGA	GCT	AAG	ATT	GGT	GCG	GGG	TCG			855
V	V	L	K	D	V	P	P	R	T	T	A	V	G	N			300
GTG	GTG	TTG	AAA	GAC	GTG	CCG	CCG	CGT	ACG	ACG	GCT	GTT	GGA	AAT			900
P	A	R	L	L	G	G	K	D	N	P	K	T	H	D			315
CCG	GCG	AGG	TTG	CTT	GGT	GGT	AAA	GAT	AAT	CCG	AAA	ACG	CAT	GAC			945
K	I	P	G	L	T	M	D	Q	T	S	H	I	S	E			330
AAG	ATT	CCT	GGT	TTG	ACT	ATG	GAC	CAG	ACG	TCG	CAT	ATA	TCC	GAG			990
W	S	D	Y	V	I												336
TGG	TCG	GAT	TAT	GTA	ATT	TGA	aaaagtc										1011
acc	agc	cta	ata	cat	taa	act	ttttcagctcgcgagagagag										
Sac 1																	

Figure 6: Séquence nucléotidique et peptidique d'un gène de l'isoforme SAT 1 (L78443) d'*A. thaliana*.

															BamH1		10							
															gagagaggatcc									
															M	L		P	V	T	S	R	R	H
ggccgagaa	aaaaaaaaa	ATG	TTG	CCG	GTC	ACA	AGT	CGC	CGC	CAC	TTC						30							
T	M	S	L	Y	M	L	R	S	S	S	P	H	I	N			25							
ACA	ATG	TCC	CTA	TAT	ATG	CTC	CGT	TCA	TCT	TCT	CCA	CAC	ATC	AAT			75							
H	H	S	F	L	L	P	S	F	V	S	S	K	F	K			40							
CAT	CAC	TCT	TTC	CTT	CTT	CCT	TCT	TTT	GTT	TCC	TCC	AAA	TTC	AAA			120							
H	H	T	L	S	P	P	P	S	P	P	P	P	P	P			55							
CAC	CAT	ACT	TTA	TCT	CCT	CCT	CCT	TCT	CCT	CCT	CCT	CCT	CCT	CCT			165							
M	A	A	C	I	D	T	C	R	T	G	K	P	Q	I			70							
ATG	GCT	GCG	TGC	ATC	GAC	ACC	TGC	CGC	ACT	GGT	AAA	CCC	CAG	ATT			210							
S	P	R	D	S	S	K	H	H	D	D	E	S	G	F			85							
TCT	CCT	CGC	GAT	TCT	TCT	AAA	CAC	CAC	GAC	GAT	GAA	TCT	GGC	TTT			255							
R	Y	M	N	Y	F	R	Y	P	D	R	S	S	F	N			100							
CGT	TAC	ATG	AAC	TAC	TTC	CGT	TAT	CCT	GAT	CGA	TCT	TCC	TTC	AAT			300							
G	T	Q	T	K	T	L	H	T	R	P	L	L	E	D			115							
GGA	ACC	CAG	ACC	AAA	ACC	CTC	CAT	ACT	CGT	CCT	TTG	CTT	GAA	GAT			345							
L	D	R	D	A	E	V	D	D	V	W	A	K	I	R			130							
CTC	GAT	CGC	GAC	GCT	GAA	GTC	GAT	GAT	GTT	TGG	GCC	AAA	ATC	CGA			390							
E	E	A	K	S	D	I	A	K	E	P	I	V	S	A			145							
GAA	GAG	GCT	AAA	TCT	GAT	ATC	GCC	AAA	GAA	CCT	ATT	GTT	TCC	GCT			435							
Y	Y	H	A	S	I	V	S	Q	R	S	L	E	A	A			160							
TAT	TAT	CAC	GCT	TCG	ATT	GTT	TCT	CAG	CGT	TCG	TTG	GAA	GCT	GCG			480							
L	A	N	T	L	S	V	K	L	S	N	L	N	L	P			175							
TTG	GCG	AAT	ACT	TTA	TCT	GTT	AAA	CTC	AGC	AAT	TTG	AAT	CTT	CCA			525							
S	N	T	L	F	D	L	F	S	G	V	L	Q	G	N			190							
AGC	AAC	ACG	CTT	TTC	GAT	TTG	TTC	TCT	GGT	GTT	CTT	CAA	GGA	AAC			570							
P	D	I	V	E	S	V	K	L	D	L	L	A	V	K			205							
CCA	GAT	ATT	GTT	GAA	TCT	GTC	AAG	CTA	GAT	CTT	TTA	GCT	GTT	AAG			615							
E	R	D	P	A	C	I	S	Y	V	H	C	F	L	H			220							
GAG	AGA	GAT	CCT	GCT	TGT	ATA	AGC	TAC	GTT	CAT	TGT	TTC	CTT	CAC			660							
F	K	G	F	L	A	C	Q	A	H	R	I	A	H	E			235							
TTT	AAA	GGC	TTC	CTC	GCT	TGT	CAA	GCG	CAT	CGT	ATT	GCT	CAT	GAG			705							
L	W	T	Q	D	R	K	I	L	A	L	L	I	Q	N			250							
CTT	TGG	ACT	CAG	GAC	AGA	AAA	ATC	CTA	GCT	TTG	TTG	ATC	CAG	AAC			750							
R	V	S	E	A	F	A	V	D	F	H	P	G	A	K			265							
AGA	GTC	TCT	GAA	GCC	TTC	GCT	GTT	GAT	TTC	CAC	CCT	GGA	GCT	AAA			795							
I	G	T	G	I	L	L	D	H	A	T	A	I	V	I			280							
ATC	GGT	ACC	GGG	ATT	TTG	CTA	GAC	CAT	GCT	ACG	GCT	ATT	GTG	ATC			840							
G	E	T	A	V	V	G	N	N	V	S	I	L	H	N			295							
GGT	GAG	ACG	GCG	GTT	GTG	GGG	AAC	AAT	GTT	TCG	ATT	CTC	CAT	AAC			885							
V	T	L	G	G	T	G	K	Q	C	G	D	R	H	P			310							
GTT	ACG	CTT	GGA	GGA	ACG	GGG	AAA	CAG	TGT	GGA	GAT	AGG	CAC	CCG			930							
K	I	G	D	G	V	L	I	G	A	G	T	C	I	L			325							
AAG	ATT	GGC	GAT	GGG	GTT	TTG	ATT	GGA	GCT	GGG	ACT	TGT	ATT	TTG			975							
G	N	I	T	I	G	E	G	A	K	I	G	A	G	S			340							
GGG	AAT	ATC	ACG	ATT	GGT	GAA	GGA	GCT	AAG	ATT	GGT	GCG	GGG	TCG			1020							
V	V	L	K	D	V	P	P	R	T	T	A	V	G	N			355							
GTG	GTG	TTG	AAA	GAC	GTG	CCG	CCG	CGT	ACG	ACG	GCT	GTT	GGA	AAT			1065							
P	A	R	L	L	G	G	K	D	N	P	K	T	H	D			370							
CCG	GCG	AGG	TTG	CTT	GGT	GGT	AAA	GAT	AAT	CCG	AAA	ACG	CAT	GAC			1110							
K	I	P	G	L	T	M	D	Q	T	S	H	I	S	E			385							
AAG	ATT	CCT	GGT	TTG	ACT	ATG	GAC	CAG	ACG	TCG	CAT	ATA	TCC	GAG			1155							
W	S	D	Y	V	I												391							
TGG	TCG	GAT	TAT	GTA	ATT	TGA	aaaagtc										1176							
acc	agc	cta	ata	cat	taa	act	ttttcagctcgcgagagag																	

Figure 7 : Séquence nucléotidique et peptidique d'un gène de l'isoforme SAT 1 (U22964) d'*A. thaliana*